

STIC-Biotech/ChemLib

178033

From: Whiteman, Brian
Sent: Monday, January 30, 2006 3:14 PM
To: STIC-Biotech/ChemLib
Subject: seq search

10/069386

SEQ ID NOs: 1 and 2

1) interference search

2) us patent and published us patent application databases

Thank you,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

RECEIVED
JAN 30 2006
STIC

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 10:33:11 ; Search time 3205 Seconds
(without alignments)
10481.897 Million cell updates/sec

Title: US-10-069-386A-1

Perfect score: 591

Sequence: 1 atggaggagagcttgaagag.....aatcatctctggggtcctaa 591

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	591	100.0	591	6	AX088925	AX088925 Sequence
2	591	100.0	1542	8	AF192529	AF192529 Homo sapi
3	589.4	99.7	1325	6	BD127913	BD127913 Primer fo
4	589.4	99.7	1325	6	CQ783944	CQ783944 Sequence
5	589.4	99.7	1325	8	AK074604	AK074604 Homo sapi
6	589.4	99.7	1346	8	BC014061	BC014061 Homo sapi
7	589.4	99.7	1420	8	BC050643	BC050643 Homo sapi
8	589.4	99.7	1536	6	CQ731553	CQ731553 Sequence
9	589.4	99.7	16063	8	AC010271	AC010271 Homo sapi
10	588.4	99.6	825	6	BD124690	BD124690 Primer fo
11	588.4	99.6	825	6	BD126646	BD126646 Primer fo
12	588.4	99.6	825	6	CQ779981	CQ779981 Sequence
13	588.4	99.6	825	6	CQ781937	CQ781937 Sequence
14	584.6	98.9	164652	14	AC021625	AC021625 Homo sapi
15	473	80.0	506	6	CQ463286	CQ463286 Sequence
16	429.2	72.6	1338	9	BC091229	BC091229 Rattus no
17	429.2	72.6	256418	14	AC115309	AC115309 Rattus no
18	429.2	72.6	257945	14	AC112313	AC112313 Rattus no

19	429.2	72.6	290929	14	AC118914	AC118914 Rattus no
20	426	72.1	657	9	AF317202	AF317202 Mus muscu
21	426	72.1	1356	9	BC034886	BC034886 Mus muscu
22	426	72.1	145347	14	AC073821	AC073821 Mus muscu
23	426	72.1	164127	14	AC158304	AC158304 Mus muscu
24	426	72.1	201377	14	AC073767	AC073767 Mus muscu
25	186.6	31.6	205	6	CQ463461	CQ463461 Sequence
26	54.6	9.2	148851	8	HS155D22	297205 Human DNA s
27	49.4	8.4	68050	14	AC079074	AC079074 Homo sapi
28	49.2	8.3	7218	6	I66494	I66494 Sequence 14
29	48.6	8.2	125020	8	AF429315	AF429315 Homo sapi
30	48	8.1	92564	8	AY007685	AY007685 Homo sapi
31	48	8.1	125020	8	AF429315	AF429315 Homo sapi
32	47.4	8.0	2082	6	AR648478	AR648478 Sequence
33	47.4	8.0	2082	9	AF322238	AF322238 Mus muscu
34	47.4	8.0	2432	9	BC055824	BC055824 Mus muscu
35	47.4	8.0	154037	9	AC126039	AC126039 Mus muscu
36	47.4	8.0	163333	9	AC160929	AC160929 Mus muscu
37	46.4	7.9	3179	6	AX003139	AX003139 Sequence
38	46.4	7.9	25138	8	HSTERT2	AF128894 Homo sapi
39	46.4	7.9	51552	6	AR266023	AR266023 Sequence
40	46.4	7.9	202305	8	AC114291	AC114291 Homo sapi
41	46	7.8	173910	8	AC083982	AC083982 Homo sapi
42	46	7.8	302176	14	AC096479	AC096479 Rattus no
43	45.6	7.7	37405	14	AC101273	AC101273 Mus muscu
44	45.6	7.7	148213	9	AL645731	AL645731 Mouse DNA
45	45.6	7.7	189378	14	AC148826	AC148826 Pan trogl

ALIGNMENTS

RESULT 1	AX088925	591 bp	DNA	linear	PAT 17-MAR-2001
LOCUS	AX088925	Sequence 1 from Patent WO0114546.			
DEFINITION	AX088925				
ACCESSION	AX088925				
VERSION	AX088925.1	GI:13397684			
KEYWORDS					
SOURCE		unidentified			
ORGANISM		unclassified			
REFERENCE	1	unclassified sequences.			
AUTHORS	Alaoui-Jamali,M.A. and Cho,J.M.				
TITLE	Replication protein a binding transcriptional factor (rbt1) and uses thereof				
JOURNAL	Patent: WO 0114546-A 1 01-MAR-2001;				
FEATURES	Centre for Translational Research in Cancer (CA)				
source	location/Qualifiers				
	1..591				
	/organism="unidentified"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32644"				
	/note="replication protein A transcriptional factor"				
ORIGIN					
Query Match	100.0%;	Score 591;	DB 6;	Length 591;	
Best Local Similarity	100.0%;	Pred. No. 9.2e-114;			
Matches 591;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	ATGAGGGAAGGCTTGAAGAGAAACACTGATTGGAAGAGAGAGAGTGGAG	60		
DB	1	ATGAGGGAAGGCTTGAAGAGAAACACTGATTGGAAGAGAGAGAGTGGAG	60		
QY	61	TGAGATCCAGAGGCTTCAAGCTACCAAGCCCTGCTCCGATCTCCAGACAA	120		
DB	61	TGAGATCCAGAGGCTTCAAGCTACCAAGCCCTGCTCCGATCTCCAGACAA	120		
QY	121	GTCAGCGAGCGCTGGGCCCCGAGACCCAGCCTCCGAGCATGCTCATCAAC	180		
DB	121	GTCAGCGAGCGCTGGGCCCCGAGACCCAGCCTCCGAGCATGCTCATCAAC	180		
QY	181	ACCCTCCAAGCTGAGGCTGCACTTGCTGCTGCCGCTGCCGCTGCCGCTGCCG	240		

|||||
Db 181 ACCCTCCAACAGCTGCAGGCTGCACTTCGCTGCTCCGCCGCCCTGCCCCGAG 240
QY 241 CCCCTCTTCTGGCGAGAGATTCTCCCTGTCAGCCACCATTTGGCTTATCTCAGG 300
Db 241 CCCCTCTTCTGGCGAGAGATTCTCCCTGTCAGCCACCATTTGGCTTATCTCAGG 300
QY 301 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTCAGATCCAGTACTCCCTGGC 360
Db 301 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTCAGATCCAGTACTCCCTGGC 360
QY 361 CTCAGATGAAGTGCACCCCAAGCCTGATCCAGTCTTTAGAAGCTTGAGCTCCCG 420
Db 361 CTCAGATGAAGTGCACCCCAAGCCTGATCCAGTCTTTAGAAGCTTGAGCTCCCG 420
QY 421 TACTTGGGGACTCTGGCCTTGATGACTTCTTTTGACATTTGACATCTGCGTAGAA 480
Db 421 TACTTGGGGACTCTGGCCTTGATGACTTCTTTTGACATTTGACATCTGCGTAGAA 480
QY 481 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCAGTTCT 540
Db 481 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCAGTTCT 540
QY 541 TGGGAGTGAATGAATGATCATCATGAAATCATTTGGGGTCTTAA 591
Db 541 TGGGAGTGAATGAATGATCATCATGAAATCATTTGGGGTCTTAA 591

RESULT 2
AF192529 1542 bp mRNA linear PRI 26-OCT-2000
LOCUS Homo sapiens RPA-binding trans-activator (RBT1) mRNA, complete cds.
DEFINITION AF192529
ACCESSION AF192529.1 GI:6180044
VERSION AF192529.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 1542)
AUTHORS Cho,J.M., Song,D.J., Bergeron,J., Benlilame,N., Wold,M.S. and Alaoui-Jamali,M.A.
TITLE RBT1, a novel transcriptional co-activator, binds the second subunit of replication protein A
JOURNAL Nucleic Acids Res. 28 (18), 3478-3485 (2000)
PUBMED 10982866
REFERENCE 2 (bases 1 to 1542)
AUTHORS Cho,J.M., Song,D.J. and Alaoui-Jamali,M.A.
TITLE Direct Submission
JOURNAL Submitted (06-OCT-1999) Experimental Medicine, Lady Davis Institute, 3999 Cote-Ste-Catherine, Montreal, QC H3T-1E2, Canada
FEATURES
source location/Qualifiers
1..1542
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1542
/gene="RBT1"
292..882
/gene="RBT1"
/codon_start=1
/product="RPA-binding trans-activator"
/protein_id="AAF05761.1"
/db_xref="GI:6180045"

Query Match 100.0%; Score 591; DB 8; Length 1542;
Best Local Similarity 100.0%; Pred. No. 8.2e-114;

Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGGAGAGCTTGAAGAGAAACACTTGATTGGAAGAGAGAGAGAGTGGAG 60
Db 292 ATGAGGAGAGCTTGAAGAGAAACACTTGATTGGAAGAGAGAGAGTGGAG 351
QY 61 TGGAGTCCAGCAGGCTTCAGAGCTACCAGCAAGCCCTGCTCCGATCTCCCTAGCAAA 120
Db 352 TGGAGTCCAGCAGGCTTCAGAGCTACCAGCAAGCCCTGCTCCGATCTCCCTAGCAAA 411
QY 121 GTCCAGCGCAGCTGGGCCCCCGAGCACCAGCCTCCGACATGTCTCATCCATAAC 180
Db 412 GTCCAGCGCAGCTGGGCCCCCGAGCACCAGCCTCCGACATGTCTCATCCATAAC 471
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTGCCTGCTCCGCCCTGCCCCCGAG 240
Db 472 ACCCTCCAACAGCTGCAGGCTGCACTTGCCTGCTCCGCCCTGCCCCCGAG 531
QY 241 CCCCTCTTCTGGCGAGAGATTCTCCCTGTCAGCCACCATTTGGCTTATCTCAGG 300
Db 532 CCCCTCTTCTGGCGAGAGATTCTCCCTGTCAGCCACCATTTGGCTTATCTCAGG 591
QY 301 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTGAATCCCTGGC 360
Db 592 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTGAATCCCTGGC 651
QY 361 CTCAGATGAAGTGCCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTTGAGCTCCCG 420
Db 652 CTCAGATGAAGTGCCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTTGAGCTCCCG 711
QY 421 TACTTGGGGACTCTGGCCTTGATGACTTCTTTTGACATTTGACATCTGCGTAGAA 480
Db 712 TACTTGGGGACTCTGGCCTTGATGACTTCTTTTGACATTTGACATCTGCGTAGAA 771
QY 481 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCAGTTCT 540
Db 772 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCAGTTCT 831
QY 541 TGGGAGTGAATGAATGATCATCATGAAATCATTTGGGGTCTTAA 591
Db 832 TGGGAGTGAATGAATGATCATCATGAAATCATTTGGGGTCTTAA 882

RESULT 3
BD127913 1325 bp DNA linear PAT 18-SEP-2002
LOCUS BD127913
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127913
VERSION BD127913.1 GI:23222858
KEYWORDS JP 2002017375-A/3344.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1325)
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Wakematsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 3344 22-JAN-2002;
COMMENT
OS Homo sapiens (human)
PN JP 2002017375-A/3344
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
PI SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/

10',
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT CDS (93)..(680).
source
1.1325
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.7%; Score 589.4; DB 6; length 1325;
Best Local Similarity 99.8%; Pred. No. 1.8e-113;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGGAGGAGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 60
Db 93 ATGTGGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 152
QY 61 TGGAGTCCAGCAGGCGCTTCAAGCTAACGACGAGCCCTGCTCCGATCTCCCTAGACAAA 120
Db 153 TGGAGTCCAGCAGGCGCTTCAAGCTAACGACGAGCCCTGCTCCGATCTCCCTAGACAAA 212
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGCATGTCTCATCCATAAC 180
Db 213 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGCATGTCTCATCCATAAC 272
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTGGCTGGCTCCCGCCCCCTGCCCCCGAG 240
Db 273 ACCCTCCAACAGCTGCAGGCTGCACTTGGCTGGCTCCCGCCCCCTGCCCCCGAG 332
QY 241 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTACGCCACCATTTGGCTATCTCTAGG 300
Db 333 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTACGCCACCATTTGGCTATCTCTAGG 392
QY 301 GAGCTGACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 393 GAGCTGACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 452
QY 361 CTCAGAATGAAGTGCCACCCCAAGCCTGATCCAGTCTTTAGAACTCTGAGTCCCG 420
Db 453 CTCAGAATGAAGTGCCACCCCAAGCCTGATCCAGTCTTTAGAACTCTGAGTCCCG 512
QY 421 TACTTGGGGGACTCTGGCCTGATGACTTCTTTTGACATTGACACATCTGCGTAGAA 480
Db 513 TACTTGGGGGACTCTGGCCTGATGACTTCTTTTGACATTGACACATCTGCGTAGAA 572
QY 481 AAGGAGCTGCACGGGGCCCCACAGAGCCTCTCAACAACCTTTCTGTGCCCCAGTTCT 540
Db 573 AAGGAGCTGCACGGGGCCCCACAGAGCCTCTCAACAACCTTTCTGTGCCCCAGTTCT 632
QY 541 TGGAGTGAATGAATGATGATCATCATGAAATCATTTGGGGTCTCTAA 591
Db 633 TGGAGTGAATGAATGATGATCATCATGAAATCATTTCTGGGGTCTCTAA 683
RESULT 4
LOCUS CQ783944 1325 bp DNA linear PAT 17-MAR-2004
DEFINITION Sequence 4084 from Patent EP1396543.
ACCESSION CQ783944
VERSION CQ783944.1 GI:45503845
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 Oca,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
AUTHORS Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.

TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 4084 10-MAR-2004;
Research Association for Biotechnology (JP)
FEATURES
source
1.1325
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
93..683
/notes="unamed protein product"
/codon_start=1
/protein_id="CAF86937.1"
/db_xref="GI:45503846"
/translation="MTVGLKRRKHSDEEBEERWSPAGLSYQALRLISLDKYQRS
LGRAPSLRRHYLIHNTLQQLAALRLAPALPEPLFGEEDFSLSATIGSTIREL
DTSMGTEPPQNPVTPGLQNEVPPODPVFLEALSSRYLGDSDLDFLDIDTSAVE
KEPARAPPEPPHNLFCAPGSWEMNELDHIMEITLGS"
CDS
ORIGIN
Query Match 99.7%; Score 589.4; DB 6; length 1325;
Best Local Similarity 99.8%; Pred. No. 1.8e-113;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGGAGGAGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 60
Db 93 ATGTGGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 152
QY 61 TGGAGTCCAGCAGGCGCTTCAAGCTAACGACGAGCCCTGCTCCGATCTCCCTAGACAAA 120
Db 153 TGGAGTCCAGCAGGCGCTTCAAGCTAACGACGAGCCCTGCTCCGATCTCCCTAGACAAA 212
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGCATGTCTCATCCATAAC 180
Db 213 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGCATGTCTCATCCATAAC 272
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTGGCTGGCTCCCGCCCCCTGCCCCCGAG 240
Db 273 ACCCTCCAACAGCTGCAGGCTGCACTTGGCTGGCTCCCGCCCCCTGCCCCCGAG 332
QY 241 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTACGCCACCATTTGGCTATCTCTAGG 300
Db 333 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTACGCCACCATTTGGCTATCTCTAGG 392
QY 301 GAGCTGACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 393 GAGCTGACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 452
QY 421 TACTTGGGGGACTCTGGCCTGATGACTTCTTTTGACATTGACACATCTGCGTAGAA 480
Db 513 TACTTGGGGGACTCTGGCCTGATGACTTCTTTTGACATTGACACATCTGCGTAGAA 572
QY 481 AAGGAGCTGCACGGGGCCCCACAGAGCCTCTCAACAACCTTTCTGTGCCCCAGTTCT 540
Db 573 AAGGAGCTGCACGGGGCCCCACAGAGCCTCTCAACAACCTTTCTGTGCCCCAGTTCT 632
QY 541 TGGAGTGAATGAATGATGATCATCATGAAATCATTTGGGGTCTCTAA 591
Db 633 TGGAGTGAATGAATGATGATCATCATGAAATCATTTCTGGGGTCTCTAA 683
RESULT 5
LOCUS AK074604 1325 bp mRNA linear PRI 03-SEP-2002
DEFINITION Homo sapiens cDNA FLJ90123 fis, clone HEMBA1007226, highly similar
to Homo sapiens RPA-binding trans-activator (RBT1) mRNA.
ACCESSION AK074604
VERSION AK074604.1 GI:22760150
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1
AUTHORS Isogai,T., Ota,T., Nishikawa,T., Hayaeshi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1325)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
Location/Qualifiers
1. 1325
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEMBA1007226"
/tissue_type="whole embryo, mainly head"
/clone_id="HEMBA1"
/dev_stage="embryo, 10 weeks"
/note="cloning vector: PME18SFL3"

ORIGIN

Query Match 99.7%; Score 589.4; DB 8; Length 1325;
Best Local Similarity 99.8%; Pred. No. 1.8e-113;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGAGGTTGGAG 60
Db 93 ATGGTGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGAGGTTGGAG 152
QY 61 TGGAGTCCAGCAGGCGCTTCAGAGCTACAGCAAGCGCTGCTCCGATCTCCCTAGACAAA 120
Db 153 TGGAGTCCAGCAGGCGCTTCAGAGCTACAGCAAGCGCTGCTCCGATCTCCCTAGACAAA 212
QY 121 GTCCAGCGCAGGCTGGGCCCCCGAGCAACCCAGCCTCCGAGGATGTCTCATCATTAAC 180
Db 213 GTCCAGCGCAGGCTGGGCCCCCGAGCAACCCAGCCTCCGAGGATGTCTCATCATTAAC 272
QY 181 ACCCTCCAACAGCTGCAGGCTGCATTGCGCTGCGTCCCGCCCCCTGCCCCCGAG 240
Db 273 ACCCTCCAACAGCTGCAGGCTGCATTGCGCTGCGTCCCGCCCCCTGCCCCCGAG 332
QY 241 CCCCTCTTCTGTTGGGCGAGGAGATTTCCTCTGTACGCCACCATTTGGCTATCTCAAG 300
Db 333 CCCCTCTTCTGTTGGGCGAGGAGATTTCCTCTGTACGCCACCATTTGGCTATCTCAAG 392
QY 301 GAGCTGACACCTTCATGATGGAGTGAAGCCCCCTCAGATTCAGTACTCCCTTGGC 360
Db 393 GAGCTGACACCTTCATGATGGAGTGAAGCCCCCTCAGATTCAGTACTCCCTTGGC 452
QY 361 CTCAGAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG 420
Db 453 CTCAGAATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTCTGAGTCCCGG 512
QY 421 TACTTTGGGGGACTGTGGCTTGATGACTTTTCTTTGTGACATTTGACATCTGCGGTAGAA 480
Db 513 TACTTTGGGGGACTGTGGCTTGATGACTTTTCTTTGTGACATTTGACATCTGCGGTAGAA 572

QY 481 AAGGAGCCTGCACGGGCCCCCACCAGAGCCTCCTCACAACTTCTTGTGCCCCCAGGTTCT 540
Db 573 AAGGAGCCTGCACGGGCCCCCACCAGAGCCTCCTCACAACTTCTTGTGCCCCCAGGTTCT 632

QY 541 TGGAGTGAATGAAGTGGATGCATCATGTGAATCATTTCTGGGGTCTAA 591
Db 633 TGGAGTGAATGAAGTGGATGCATCATGTGAATCATTTCTGGGGTCTAA 683

RESULT 6
BC014061 1346 bp mRNA linear PRI 25-JUL-2005
LOCUS
DEFINITION Homo sapiens SERTA domain containing 3, transcript variant 2, mRNA
(cDNA clone MGC:19994 IMAGE:3926937), complete cds.
ACCESSION BC014061
VERSION BC014061.2 GI:33873834
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1346)
AUTHORS Krausberg,R.L., Felngold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Muliahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Young,A.C., Shevchenko,Y.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Mammalian Gene Collection Program Team
JOURNAL Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1346)

CONSRM
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
CONSRM
TITLE
JOURNAL
REMARK
COMMENT
NIH MGC Project
Direct Submission
Submitted (10-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
On Aug 19, 2003 this sequence version replaced gi:15559388.
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nigri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakeley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Grantle,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Lalic,P., Legaspi,R.,
Maduro,Q.L., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantriop,S., Thomas,P.J., Touchman,J.W.,
Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAL Plate: 28 Row: f Column: 10
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 15718683.

FEATURES

source

1.1346
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:19994 IMAGE:3926937"
/issue_type="Ovary, adenocarcinoma"
/clone_lib="NIH_MGC_9"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1.1346
/gene="SERTAD3"
/note="synonym: RBT1"
/db_xref="GeneID:29946"
92.682
/gene="SERTAD3"
/codon_start=1
/product="RPA-binding trans-activator"
/protein_id="AAH14061.1"
/db_xref="GI:15559389"
/db_xref="GeneID:29946"
/translation="MWGLKRRKHSDBBBBERWESPAGLSYQQALRLISLDKVOBS
LGRAPSLRRHVLJHNTLQOLQALRLAPALPPEPLFEGEDFSLSATIGSLREL
DTSMDTEPPQNPVTPGLQNEVPPQDPVPLFALSRYLGDSDLDFLIDITSAYE
KEPARAPEPRPHNLFCAPSWEMNELDHIMEILIGS"

ORIGIN

Query Match 99.7%; Score 589.4; DB 8; Length 1346;
Best Local Similarity 99.8%; Pred. No. 1.8e-113;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGGAGGAGGCTTGAAGAGAAACACTGTGATTTGAAGAGAGAGAGAGAGGTGGAG 60
Db 92 ATGGTGGAGGCTTGAAGAGAAACACTGTGATTTGAAGAGAGAGAGAGAGGTGGAG 151
OY 61 TGGAGTCCAGCAGGCGCTTCAGAGCTAACAGCAGGCCCTGCTCCGCATCTCCCTAGACAA 120
Db 152 TGGAGTCCAGCAGGCGCTTCAGAGCTAACAGCAGGCCCTGCTCCGCATCTCCCTAGACAA 211
OY 121 GTCCAGCGCAGCCTGGGCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCACTAAC 180
Db 212 GTCCAGCGCAGCCTGGGCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCACTAAC 271
OY 181 ACCCTCCAACAGCTGCAGGCTGCACTTGCGCTGCTCCGCCCTGCGCTGCCCGCGAG 240
Db 272 ACCCTCCAACAGCTGCAGGCTGCACTTGCGCTGCTCCGCCCTGCGCTGCCCGCGAG 331
OY 241 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTACGCCACCATTTGGCTCTATCCTCAGG 300
Db 332 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTACGCCACCATTTGGCTCTATCCTCAGG 391
OY 301 GAGCTGCACACCTCCATGATGGGAAGTGAAGCCCCCTCAGAATCCAGTACTCCCCCTTGGC 360
Db 392 GAGCTGCACACCTCCATGATGGGAAGTGAAGCCCCCTCAGAATCCAGTACTCCCCCTTGGC 451
OY 361 CTCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTTTGAAGAGCTTGAGCTCCCGG 420
Db 452 CTCAGAAATGAAGTGCCACCCAGCCTGATCCAGTCTTTTGAAGAGCTTGAGCTCCCGG 511
OY 421 TACTTGGGGAGCTTGCGCTGAGTGAATCTTTCTTGACATTTGACACATCTGCGGTAGAA 480
Db 512 TACTTGGGGAGCTTGCGCTGAGTGAATCTTTCTTGACATTTGACACATCTGCGGTAGAA 571
OY 481 AAGGAGCTGCACGCGGCCCAACAGAGCCTCTCAACAACCTTTCTGTGCCCAAGTTCT 540
Db 572 AAGGAGCTGCACGCGGCCCAACAGAGCCTCTCTCAACAACCTTTCTGTGCCCAAGTTCT 631

OY 541 TGGAGTGAATGAAGTCAATCATCATGGAATCATTTCTGGGGTCTAA 591
Db 632 TGGAGTGAATGAAGTCAATCATCATGGAATCATTTCTGGGGTCTAA 682

RESULT 7
BC050643
LOCUS
DEFINITION
Homo sapiens SERTA domain containing 3, transcript variant 2, mRNA (cDNA clone MGC:60134 IMAGE:6503515), complete cds.
BC050643
BC050643.2 GI:34193422
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 (bases 1 to 1420)
Serausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg B, Bueltow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Heieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Vialalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Kettelman M, Madan A, Rodriguez S, Sanchez A, Whitting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skelaka U, Smailus DE, Schnerch A, Schein JE, Jones SJ and Marra MA.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1420)
Director MGC Project.

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (08-APR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 25, 2003 this sequence version replaced gi:29791864.
Contact: MGC help desk
Email: cgabs-rt@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAL Plate: 110 Row: o Column: 24
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 15718683.

FEATURES

source

1.1420
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:60134 IMAGE:6503515"
/issue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_71"


```

gene
    /lab_host="DH10B"
    /note="Vector: pCMV-SPORT6"
    1..1420
    /gene="SERTAD3"
    /note="synonym: RBT1"
    /db_xref="GeneID:29946"
    168..758
    /gene="SERTAD3"
    /codon_start=1
    /product="RPA-binding trans-activator"
    /protein_id="AAH50643.1"
    /db_xref="GI:29791865"
    /db_xref="GeneID:29946"
    /translation="MWGGLKHKHSDLEBBEREMWSPAGLSYQALLRISLDKVRSL
    LGRAPSLRRLHNTLQQLQALRLAPALPPEPLFLGEBDFSLSATIGSLREL
    DTSMDTEPPQNPVYTPGLQNEVPPQDPVPLIALSSRYLGDSGLDPLDIDTSAVE
    KEPARAPPEPPHNLFCAPGSWEMNELDHIMEIILIGS"

CDS

ORIGIN
Query Match          99.7%; Score 589.4; DB 8; Length 1420;
Best Local Similarity 99.8%; Pred. No. 1.8e-113;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGGAGGAGCTTGAAGAAGAAACACTGTGATTGTGAAGAGAGAGAGAGGTGGAG 60
    |||
Db 168 ATGTGGGAGGCTTGAAGAGAAACACTGTGATTGTGAAGAGAGAGAGAGGTGGAG 227
QY 61 TGGAGTCCAGCAGGCGCTTCAAGCTAACAGCAGAGCCCTGCTCCGCATCTCCCTAGACAA 120
    |||
Db 228 TGGAGTCCAGCAGGCGCTTCAAGCTAACAGCAGAGCCCTGCTCCGCATCTCCCTAGACAA 287
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGAGGCAATGCTCATCCATAAC 180
    |||
Db 288 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGAGGCAATGCTCATCCATAAC 347
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTGCGCTGGCTCCGCCCCCTGCCCCCGAG 240
    |||
Db 348 ACCCTCCAACAGCTGCAGGCTGCACTTGCGCTGGCTCCGCCCCCTGCCCCCGAG 407
QY 241 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTCAAGCCACCATTTGGCTATCTCAGG 300
    |||
Db 408 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTCAAGCCACCATTTGGCTATCTCAGG 467
QY 301 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCCCTGGC 360
    |||
Db 468 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCCCTGGC 527
QY 361 CTCAGAATGAAGTGCACCCCCAGCTGTATCCAAGTCTTTAGAAGCTCTGAGTCCCG 420
    |||
Db 528 CTCAGAATGAAGTGCACCCCCAGCTGTATCCAAGTCTTTAGAAGCTCTGAGTCCCG 587
QY 421 TACTTGGGGGACTCTGGGCTGGATGACTTCTTTCTGGACATTTGACACATCTGCGGTAGAA 480
    |||
Db 588 TACTTGGGGGACTCTGGGCTGGATGACTTCTTTCTGGACATTTGACACATCTGCGGTAGAA 647
QY 481 AAGGAGCTGCACGCGCCCCACAGAGCCTCTCACAACCTCTTCTGTGCCCCCAGGTTCT 540
    |||
Db 648 AAGGAGCTGCACGCGCCCCACAGAGCCTCTCACAACCTCTTCTGTGCCCCCAGGTTCT 707
QY 541 TGGGAGTGAATGAATGATCATCATCATGGAATCATTTCTGGGGTCTTAA 591
    |||
Db 708 TGGGAGTGAATGAATGATCATCATCATGGAATCATTTCTGGGGTCTTAA 758

RESULT 8
LOCUS      CQ731553          1536 bp      DNA      linear      PAT 03-FEB-2004
DEFINITION Sequence 17487 from Patent WO02068579.
ACCESSION  CQ731553
VERSION    CQ731553.1   GI:42308440
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM  Homo sapiens
```

```

REFERENCE
AUTHORS    1
TITLE      Venter,C.J., Adams,M.C., Li,P.W. and Myers,B.W.
JOURNAL    Patent: WO 02068579-A 17487 06-SEP-2002;
FEATURES
SOURCE     Location/Qualifiers
            1..1536
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match          99.7%; Score 589.4; DB 6; Length 1536;
Best Local Similarity 99.8%; Pred. No. 1.8e-113;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGGAGGAGCTTGAAGAGAAACACTGTGATTGTGAAGAAGAGAGAGAGGTGGAG 60
    |||
Db 290 ATGTGGGAGGCTTGAAGAGAAACACTGTGATTGTGAAGAAGAGAGAGAGGTGGAG 349
QY 61 TGGAGTCCAGCAGGCGCTTCAAGCTAACAGCAGAGCCCTGCTCCGCATCTCCCTAGACAA 120
    |||
Db 350 TGGAGTCCAGCAGGCGCTTCAAGCTAACAGCAGAGCCCTGCTCCGCATCTCCCTAGACAA 409
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGAGGCAATGCTCATCCATAAC 180
    |||
Db 410 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGAGGCAATGCTCATCCATAAC 469
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTGCGCTGGCTCCGCCCCCTGCCCCCGAG 240
    |||
Db 470 ACCCTCCAACAGCTGCAGGCTGCACTTGCGCTGGCTCCGCCCCCTGCCCCCGAG 529
QY 241 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTCAAGCCACCATTTGGCTATCTCAGG 300
    |||
Db 530 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTCAAGCCACCATTTGGCTATCTCAGG 589
QY 301 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCCCTGGC 360
    |||
Db 590 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCCCTGGC 649
QY 361 CTCAGAATGAAGTGCACCCCCAGCTGTATCCAAGTCTTTAGAAGCTCTGAGTCCCG 420
    |||
Db 650 CTCAGAATGAAGTGCACCCCCAGCTGTATCCAAGTCTTTAGAAGCTCTGAGTCCCG 709
QY 421 TACTTGGGGGACTCTGGGCTGGATGACTTCTTTCTGGACATTTGACACATCTGCGGTAGAA 480
    |||
Db 710 TACTTGGGGGACTCTGGGCTGGATGACTTCTTTCTGGACATTTGACACATCTGCGGTAGAA 769
QY 481 AAGGAGCTGCACGCGCCCCACAGAGCCTCTCACAACCTCTTCTGTGCCCCCAGGTTCT 540
    |||
Db 770 AAGGAGCTGCACGCGCCCCACAGAGCCTCTCACAACCTCTTCTGTGCCCCCAGGTTCT 829
QY 541 TGGGAGTGAATGAATGATCATCATCATGGAATCATTTCTGGGGTCTTAA 591
    |||
Db 830 TGGGAGTGAATGAATGATCATCATCATGGAATCATTTCTGGGGTCTTAA 880

RESULT 9
LOCUS      AC010271/c          160643 bp      DNA      linear      PRI 13-JUL-2002
DEFINITION Homo sapiens chromosome 19 clone CTC-492K19, complete sequence.
ACCESSION  AC010271
VERSION    AC010271.8   GI:21743753
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

REFERENCE 1 (bases 1 to 160643)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 160643)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 160643)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 160643)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (03-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 5 (bases 1 to 160643)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (13-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 13, 2002 this sequence version replaced gi:15887272.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.1.
FEATURES
source 1..160643
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTC-492K19"
ORIGIN
Query Match 99.7%; Score 589.4; DB 8; Length 160643;
Best Local Similarity 99.8%; Pred. No. 1e-113;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 60
Db 137453 ATGATGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 137394
QY 61 TGGAGTCCAGCAGGCTTCAGAGCTACCAAGAGCCCTGCTCCGCATCTCCCTAGACAA 120
Db 137393 TGGAGTCCAGCAGGCTTCAGAGCTACCAAGAGCCCTGCTCCGCATCTCCCTAGACAA 137334
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGCAAGCATGTCTCATCATTAAC 180
Db 137333 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGCAAGCATGTCTCATCATTAAC 137274
QY 181 ACCCTCCAACAGCTGCAAGGCTGCACTTGGCTGGCTCCCGCCCCCTGCCCCCGAG 240
Db 137273 ACCCTCCAACAGCTGCAAGGCTGCACTTGGCTGGCTCCCGCCCCCTGCCCCCGAG 137214
QY 241 CCCCTCTTCTGGGCGAGAGGATTCTCTCCCTGTGAGCCACCATTTGGCTTATCTCAGG 300
Db 137213 CCCCTCTTCTGGGCGAGAGGATTCTCTCCCTGTGAGCCACCATTTGGCTTATCTCAGG 137154
QY 301 GAGCTGACACCTCCATGATGGAAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 137153 GAGCTGACACCTCCATGATGGAAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 137094
QY 361 CTCAGAATGAAGTGCCACCCCGAGCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCGG 420
Db 137093 CTCAGAATGAAGTGCCACCCCGAGCTGATCCAGTCTTCTTAGAAGCTCTGAGCTCCGG 137034

QY 421 TACTTGGGGGACTCTGGCCTGATGACTTCTTCTTGAGCATTTGACACATCTGCGGTAGA 480
Db 137033 TACTTGGGGGACTCTGGCCTGATGACTTCTTCTTGAGCATTTGACACATCTGCGGTAGA 136974
QY 481 AAGAGCCTGCACGGGCCCCCAGACAGAGCCTCTCTCACAACCTCTTGTGCCCCAGGTCT 540
Db 136973 AAGAGCCTGCACGGGCCCCCAGACAGAGCCTCTCTCACAACCTCTTGTGCCCCAGGTCT 136914
QY 541 TGGAGTGAATGAAGTGAATGATCATCATGGAATCATTTGGGGTCTTAA 591
Db 136913 TGGAGTGAATGAAGTGAATGATCATCATGGAATCATTTGGGGTCTTAA 136863
RESULT 10
BD124690
LOCUS
DEFINITION BD124690 825 bp DNA linear PAT 18-SEP-2002
ACCESSION BD124690
VERSION BD124690.1 GI:23219635
KEYWORDS JP 2002017375-A/121.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 825)
Ota,T., Nishikawa,T., Isogai,T., Hayaeshi,K., Ishii,S., Kawai,Y., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.
Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 121 22-JAN-2002;
HEPIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/121
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10,
C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
location/Qualifiers
FT source 1..825
/organism='Homo sapiens (human)'.
FEATURES
source 1..825
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.6%; Score 588.4; DB 6; Length 825;
Best Local Similarity 99.7%; Pred. No. 3.1e-113;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGGAGGCTTGAAGAGAAACACTGTGATTGGAAGAGAGAGAGAGGTGGAG 60
Db 93 ATGATGGGAGGCTTGAAGAGAAACACTGTGATTGGAAGAGAGAGAGAGGTGGAG 152
QY 61 TGGAGTCCAGCAGGCTTCAGAGCTAACCAAGAGCCCTGCTCCGCATCTCCCTAGACAA 120
Db 153 TGGAGTCCAGCAGGCTTCAGAGCTAACCAAGAGCCCTGCTCCGCATCTCCCTAGACAA 212
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGCAAGCATGTCTCATCATTAAC 180
Db 213 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGCAAGCATGTCTCATCATTAAC 272
QY 181 ACCCTCCAACAGCTGCAAGGCTGCACTTGGCTGGCTCCCGCCCCCTGCCCCCGAG 240

|||||
Db 273 ACCCTCCACAGCTGCAGGCTGCACCTTCGCTCGCTCCGCCCTGCCCCCTGCCCCGAG 332
QY 241 CCCCTCTTCTGGCGAGAGGATTTCTCCCTGTACGCCACCATTTGGCTTATCTCAGG 300
|||||
Db 333 CCCCTCTTCTGGCGAGAGGATTTCTCCCTGTACGCCACCATTTGGCTTATCTCAGG 392
QY 301 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTGAGAATCCAGTACTCCCTTGGC 360
|||||
Db 393 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTGAGAATCCAGTACTCCCTTGGC 452
QY 361 CTCCAGAATGAAGTGCCACCCCAAGCTGTATCCAGTCTTTTGAAGCTGTAGCTCCCG 420
|||||
Db 453 CTCCAGAATGAAGTGCCACCCCAAGCTGTATCCAGTCTTTTGAAGCTGTAGCTCCCG 512
QY 421 TACTTGGGGGACTCTGGCTGTGATGACTTCTTTCTGACATTTGACACATCTGCGGTAGAA 480
Db 513 TACTTGGGGGACTCTGGCTGTGATGACTTCTTTCTGACATTTGACACATCTGCGGTAGAA 572
QY 481 AAGAGCCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGTGCCCCAGTTCT 540
Db 573 AAGAGCCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGTGCCCCAGTTCT 632
QY 541 TGGAGTGAATGAATGATCATCATCATGAAATCATTTCTGGGTCCTAA 591
|||||
Db 633 TGGAGTGAATGAATGATCATCATCATGAAATCATTTCTGGGTCCTAA 683

RESULT 11
BD126646 825 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD126646
VERSION BD126646.1 GI:23221591
KEYWORDS JP 2002017375-A/2077.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 825)
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 2077 22-JAN-2002;
COMMENT
HELEX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/2077
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
FT source 1..825
FT location/Qualifiers
FT /organism='Homo sapiens (human)'.
FT 1..825
FT /organism="Homo sapiens"
FT /mol_type="genomic DNA"
FT /db_xref="taxon:9606"
ORIGIN
Query Match 99.6%; Score 588.4; DB 6; Length 825;
Best Local Similarity 99.7%; Pred. No. 3.1e-113;

Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGGAGGCTTGAAGAGAAACACTGTATTTGGAAGAGAGAGAGGTGGAG 60
|||||
Db 93 ATGTTGGAGGCTTGAAGAGAAACACTGTATTTGGAAGAGAGAGAGGTGGAG 152
QY 61 TGGAGTCCAGCAGGCTTCAAGACTACAGCAAGCCTGTCTCCGATCTCCCTAGCAAA 120
|||||
Db 153 TGGAGTCCAGCAGGCTTCAAGACTACAGCAAGCCTGTCTCCGATCTCCCTAGCAAA 212
QY 121 GTCCAGCGCAGCTTGGGCCCCCGAGCACCAGCCTCCGCAAGCATGTCTCATCCATAAC 180
Db 213 GTCCAGCGCAGCTTGGGCCCCCGAGCACCAGCCTCCGCAAGCATGTCTCATCCATAAC 272
QY 181 ACCCTCCACAGCTGCAGGCTGCACCTTGCCTGTCCGCCCTGCCCCCGAG 240
Db 273 ACCCTCCACAGCTGCAGGCTGCACCTTGCCTGTCCGCCCTGCCCCCGAG 332
QY 241 CCCCTCTTCTGGCGAGAGGATTTCTCCCTGTACGCCACCATTTGGCTTATCTCAGG 300
Db 333 CCCCTCTTCTGGCGAGAGGATTTCTCCCTGTACGCCACCATTTGGCTTATCTCAGG 392
QY 301 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTGAGAATCCAGTACTCCCTTGGC 360
Db 393 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTGAGAATCCAGTACTCCCTTGGC 452
QY 361 CTCCAGAATGAAGTGCCACCCCAAGCCTGATCCAGTCTTTTGAAGCTCTGAGCTCCCG 420
Db 453 CTCCAGAATGAAGTGCCACCCCAAGCCTGATCCAGTCTTTTGAAGCTCTGAGCTCCCG 512
QY 421 TACTTGGGGGACTCTGGCTGTGATGACTTCTTTCTGACATTTGACACATCTGCGGTAGAA 480
Db 513 TACTTGGGGGACTCTGGCTGTGATGACTTCTTTCTGACATTTGACACATCTGCGGTAGAA 572
QY 481 AAGAGCCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGTGCCCCAGTTCT 540
Db 573 AAGAGCCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGTGCCCCAGTTCT 632
QY 541 TGGAGTGAATGAATGATCATCATCATGAAATCATTTCTGGGTCCTAA 591
|||||
Db 633 TGGAGTGAATGAATGATCATCATCATGAAATCATTTCTGGGTCCTAA 683

RESULT 12
CQ779981 825 bp DNA linear PAT 17-MAR-2004
LOCUS CQ779981
DEFINITION Sequence 121 from Patent EP1396543.
ACCESSION CQ779981
VERSION CQ779981.1 GI:45536053
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 121 10-MAR-2004;
COMMENT Research Association for Biotechnology (JP)
FEATURES
FT source 1..825
FT location/Qualifiers
FT /organism="Homo sapiens"
FT /mol_type="unassigned DNA"
FT /db_xref="taxon:9606"
ORIGIN

Query Match 99.6%; Score 588.4; DB 6; Length 825;
Best Local Similarity 99.7%; Pred. No. 3.1e-113;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGGAGGCTTGAAGAGAAACACTGTGATTGGAGAGAGAGAGAGAGGTGGAG 60
|||||
Db 93 ATGTGGGAGGCTTGAAGAGAAACACTGTGATTGGAGAGAGAGAGAGAGGTGGAG 152
QY 61 TGGAGTCCAGCAGGCGCTTCAGAGCTAACCAAGACCCCTGCTCCGCATCTCCCTAGACAA 120
|||||
Db 153 TGGAGTCCAGCAGGCGCTTCAGAGCTAACCAAGACCCCTGCTCCGCATCTCCCTAGACAA 212
QY 121 GTCCAGCGCAGCCTGGGCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC 180
|||||
Db 213 GTCCAGCGCAGCCTGGGCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC 272
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTCGCTGCTCCGCCCCCTGCCCCGCGAG 240
|||||
Db 273 ACCCTCCAACAGCTGCAGGCTGCACTTCGCTGCTCCGCCCCCTGCCCCGCGAG 332
QY 241 CCCCTCTTCCTGGGCGAGAGGATTCTCCCTGTCAAGCCACCATTTGGCTCTATCCTCAGG 300
|||||
Db 333 CCCCTCTTCCTGGGCGAGAGGATTCTCCCTGTCAAGCCACCATTTGGCTCTATCCTCAGG 392
QY 301 GAGCTGGACACCTCCATGATGGGAGCTGAGCCCCCTCAGAATCCAGTACTCCCCCTTGGC 360
|||||
Db 393 GAGCTGGACACCTCCATGATGGGAGCTGAGCCCCCTCAGAATCCAGTACTCCCCCTTGGC 452
QY 361 CTCAGAATGAAGTGCCACCCCAAGCCTGTAGTCCAGTCTTTAGAAAGCTCTGAGCTCCGG 420
|||||
Db 453 CTCAGAATGAAGTGCCACCCCAAGCCTGTAGTCCAGTCTTTAGAAAGCTCTGAGCTCCGG 512
QY 421 TACTTGGGGGACTCTGGCCTGGATGACTTTCTTCTGACATTTGACACATCTGCGGTAGAA 480
|||||
Db 513 TACTTGGGGGACTCTGGCCTGGATGACTTTCTTCTGACATTTGACACATCTGCGGTAGAA 572
QY 481 AAGGAGCCTGCACGCGGCCCAACAGAGCCTCCTCACAACCTCTTCTGTGCCCCAGTTCT 540
|||||
Db 573 AAGGAGCCTGCACGCGGCCCAACAGAGCCTCCTCACAACCTCTTCTGTGCCCCAGTTCT 632
QY 541 TGGGAGTGAATGAATGATGATCATCATGGAATCATTTCTGGGGTCTTAA 591
|||||
Db 633 TGGGAGTGAATGAATGATGATCATCATGGAATCATTTCTGGGGTCTTAA 683

RESULT 13
LOCUS CQ781937 825 bp DNA linear PAT 17-MAR-2004
DEFINITION Sequence 2077 from Patent EP1396543.
ACCESSION CQ781937
VERSION CQ781937.1 GI:45537993
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS 1
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: BP 1396543-A 2077 10-MAR-2004;
Research Association for Biotechnology (JP)
FEATURES
Source
1. 825
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.6%; Score 588.4; DB 6; Length 825;
Best Local Similarity 99.7%; Pred. No. 3.1e-113;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGGAGGCTTGAAGAGAAACACTGTGATTGGAGAGAGAGAGAGAGGTGGAG 60
|||||

Db 93 ATGTGGGAGGCTTGAAGAGAAACACTGTGATTGGAGAGAGAGAGAGAGGTGGAG 152
QY 61 TGGAGTCCAGCAGGCGCTTCAGAGCTAACCAAGACCCCTGCTCCGCATCTCCCTAGACAA 120
|||||
Db 153 TGGAGTCCAGCAGGCGCTTCAGAGCTAACCAAGACCCCTGCTCCGCATCTCCCTAGACAA 212
QY 121 GTCCAGCGCAGCCTGGGCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC 180
|||||
Db 213 GTCCAGCGCAGCCTGGGCCCCGAGCACCCAGCCTCCGAGGCATGTCTCATCCATAAC 272
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTCGCTGCTCCGCCCCCTGCCCCGCGAG 240
|||||
Db 273 ACCCTCCAACAGCTGCAGGCTGCACTTCGCTGCTCCGCCCCCTGCCCCGCGAG 332
QY 241 CCCCTCTTCCTGGGCGAGAGGATTCTCCCTGTCAAGCCACCATTTGGCTCTATCCTCAGG 300
|||||
Db 333 CCCCTCTTCCTGGGCGAGAGGATTCTCCCTGTCAAGCCACCATTTGGCTCTATCCTCAGG 392
QY 301 GAGCTGGACACCTCCATGATGGGAGCTGAGCCCCCTCAGAATCCAGTACTCCCCCTTGGC 360
|||||
Db 393 GAGCTGGACACCTCCATGATGGGAGCTGAGCCCCCTCAGAATCCAGTACTCCCCCTTGGC 452
QY 361 CTCAGAATGAAGTGCCACCCCAAGCCTGTAGTCCAGTCTTTAGAAAGCTCTGAGCTCCGG 420
|||||
Db 453 CTCAGAATGAAGTGCCACCCCAAGCCTGTAGTCCAGTCTTTAGAAAGCTCTGAGCTCCGG 512
QY 421 TACTTGGGGGACTCTGGCCTGGATGACTTTCTTCTGACATTTGACACATCTGCGGTAGAA 480
|||||
Db 513 TACTTGGGGGACTCTGGCCTGGATGACTTTCTTCTGACATTTGACACATCTGCGGTAGAA 572
QY 481 AAGGAGCCTGCACGCGGCCCAACAGAGCCTCCTCACAACCTCTTCTGTGCCCCAGTTCT 540
|||||
Db 573 AAGGAGCCTGCACGCGGCCCAACAGAGCCTCCTCACAACCTCTTCTGTGCCCCAGTTCT 632
QY 541 TGGGAGTGAATGAATGATGATCATCATGGAATCATTTCTGGGGTCTTAA 591
|||||
Db 633 TGGGAGTGAATGAATGATGATCATCATGGAATCATTTCTGGGGTCTTAA 683

RESULT 14
LOCUS AC021625 164652 bp DNA linear HTG 28-MAR-2000
DEFINITION Homo sapiens clone RP11-384B6, WORKING DRAFT SEQUENCE, 34 unordered
pieces.
ACCESSION AC021625
VERSION AC021625.2 GI:7331453
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 164652)
TITLE Birren,B., Linton,L., Nusbaum,C. and Lander,B.
JOURNAL Homo sapiens, clone RP11-384B6
AUTHORS Unpublished
2 (bases 1 to 164652)

Birren,B., Linton,L., Nusbaum,C., Lander,B., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewart,K., Domino,M., Doyle,M., Fenesfor,J.,
Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lander,S., Lander,T., Lehoczkzy,J., Levine,R., Liu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheters,R., Meldrim,J., Menues,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,

TITLE
JOURNAL
COMMENT

Zimmer, A. and Zody, M.

Direct Submission

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 28, 2000 this sequence version replaced gi:6705474.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L5880

Center clone name: 384_R_6

----- Summary Statistics

Sequencing vector: M13; M77815, 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 145359 bases at least Q40

Consensus quality: 155420 bases at least Q30

Consensus quality: 159351 bases at least Q20

Insert size: 161352; sum-of-contigs

Quality coverage: 4.0 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1083: contig of 1083 bp in length
* 1084 1183: gap of 100 bp
* 1184 2479: contig of 1296 bp in length
* 2480 2579: gap of 100 bp
* 2580 4311: contig of 1732 bp in length
* 4312 4411: gap of 100 bp
* 4412 5951: contig of 1540 bp in length
* 5952 6051: gap of 100 bp
* 6052 7761: contig of 1710 bp in length
* 7762 7861: gap of 100 bp
* 7862 9630: contig of 1769 bp in length
* 9631 9730: gap of 100 bp
* 9731 11763: contig of 2033 bp in length
* 11764 11863: gap of 100 bp
* 11864 14294: contig of 2431 bp in length
* 14295 14394: gap of 100 bp
* 14395 16731: contig of 2337 bp in length
* 16732 16831: gap of 100 bp
* 16832 20279: contig of 3448 bp in length
* 20280 20379: gap of 100 bp
* 20380 23751: contig of 3372 bp in length
* 23752 23851: gap of 100 bp
* 23852 26903: contig of 3052 bp in length
* 26904 27003: gap of 100 bp
* 27004 31228: contig of 4225 bp in length
* 31229 31328: gap of 100 bp
* 31329 35336: contig of 4008 bp in length
* 35337 35436: gap of 100 bp
* 35437 38469: contig of 3033 bp in length
* 38470 38569: gap of 100 bp
* 38570 42643: contig of 4074 bp in length
* 42644 42743: gap of 100 bp
* 42744 47087: contig of 4344 bp in length
* 47088 47187: gap of 100 bp
* 47188 51475: contig of 4288 bp in length
* 51476 51575: gap of 100 bp
* 51576 56093: contig of 4518 bp in length
* 56094 56194 59387: gap of 100 bp
* 56194 59387: contig of 3194 bp in length

* 59388 59487: gap of 100 bp
* 59488 63987: contig of 4500 bp in length
* 63988 64087: gap of 100 bp
* 64088 69347: contig of 5260 bp in length
* 69348 69447: gap of 100 bp
* 69448 74268: contig of 4821 bp in length
* 74269 74368: gap of 100 bp
* 74369 79365: contig of 4997 bp in length
* 79366 79465: gap of 100 bp
* 79466 83211: contig of 3746 bp in length
* 83212 83311: gap of 100 bp
* 83312 86101: contig of 4790 bp in length
* 86102 88201: gap of 100 bp
* 88102 88201: gap of 100 bp
* 88202 93785: contig of 5584 bp in length
* 93786 93885: gap of 100 bp
* 93886 100407: contig of 6522 bp in length
* 100408 100507: gap of 100 bp
* 100508 108158: contig of 7651 bp in length
* 108159 108258: gap of 100 bp
* 108259 115631: contig of 7373 bp in length
* 115632 115731: gap of 100 bp
* 115732 125488: contig of 9757 bp in length
* 125489 125588: gap of 100 bp
* 125589 136013: contig of 10425 bp in length
* 136014 136113: gap of 100 bp
* 136114 149110: contig of 12997 bp in length
* 149111 149210: gap of 100 bp
* 149211 164652: contig of 15442 bp in length.

FEATURES

source

1. .164652
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-384E6"
/clone_1fb="RPC1-11 Human Male BAC"
1. .1083
/note="assembly_fragment"
1084. .1183
/estimated_length=100
1184. .2479
/note="assembly_fragment"
2480. .2579
/estimated_length=100
2580. .4311
/note="assembly_fragment"
4312. .4411
/estimated_length=100
4412. .5951
/note="assembly_fragment"
5952. .6051
/estimated_length=100
6052. .7761
/note="assembly_fragment"
7762. .7861
/estimated_length=100
7862. .9630
/note="assembly_fragment"
9631. .9730
/estimated_length=100
9731. .11763
/note="assembly_fragment"
11764. .11863
/estimated_length=100
11864. .14294
/note="assembly_fragment"
14295. .14394
/estimated_length=100
14395. .16731
/note="assembly_fragment"
16732. .16831
/estimated_length=100
16832. .20279
/note="assembly_fragment"

gap 20280. .20379
/estimated_length=100
misc_feature 20380. .23751
/note="assembly_fragment"
gap 23752. .23851
/estimated_length=100
misc_feature 23852. .26903
/note="assembly_fragment"
gap 26904. .27003
/estimated_length=100
misc_feature 27004. .31228
/note="assembly_fragment"
gap 31229. .31328
/estimated_length=100
misc_feature 31329. .35336
/note="assembly_fragment"
gap 35337. .35436
/estimated_length=100
misc_feature 35437. .38469
/note="assembly_fragment"
gap 38470. .38569
/estimated_length=100

Query Match 98.9%; Score 584.6; DB 14; length 164652;
Best Local Similarity 99.3%; Pred. No. 1e-112;
Matches 587; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGAGGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 60
Db 113509 ATGGTGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 113568
QY 61 TGGAGTCCAGCAGGCTTCAAGCTACCAAGAGCCCTGCTCCGATCTCCCTAGACAA 120
Db 113569 TGGAGTCCAGCAGGCTTCAAGCTACCAAGAGCCCTGCTCCGATCTCCCTAGACAA 113628
QY 121 GTCCAGCGCAGCCTGGGCCCCGAGCAACCCAGCCTCCGAGCATGTCTCATCCATAAC 180
Db 113629 GTCCAGCGCAGCCTGGGCCCCGAGCAACCCAGCCTCCGAGCATGTCTCATCCATAAC 113688
QY 181 ACCCTCCAACAGCTGAGGCTGCACTTGCCCTGGCTCCGCCCCCTGCCCCCGAG 240
Db 113689 ACCCTCCAACAGCTGAGGCTGCACTTGCCCTGGCTCCGCCCCCTGCCCCCGAG 113748
QY 241 CCCCTCTTCTGGGCGAGAGGATTCTCCCTGTCAAGCCACCATTTGGCTTATCTCAGG 300
Db 113749 CCCCTCTTCTGGGCGAGAGGATTCTCCCTGTCAAGCCACCATTTGGCTTATCTCAGG 113808
QY 301 GAGCTGACACCTCCATGATGAGGAGCTGAGCCCCCTCAGAATCCAGTACTCCCTTGGC 360
Db 113809 GAGCTGACACCTCCATGATGAGGAGCTGAGCCCCCTCAGAATCCAGTACTCCCTTGGC 113868
QY 361 CTCCAGAATGAAGTGCACCCCGAGCCTGATCCAGTCTTTAGAAAGCTTGAGCTCCCG 420
Db 113869 CTCCAGAATGAAGTGCACCCCGAGCCTGATCCAGTCTTTAGAAAGCTTGAGCTCCCG 113928
QY 421 TACTTGGGGGACTCTGGCCTGATGACTTTCTTCTGACATTGACACATCTGCGGTAGAA 480
Db 113929 TACTTGGGGGACTCTGGCCTGATGACTTTCTTCTGACATTGACACATATGCGGTAGAA 113988
QY 481 AAGAGCCTGCACGGGCCCCACAGAGCCTCTCACAACCTCTTGTGCCCCCAGGTTCT 540
Db 113989 AAGAGCCTGCACGGGCCCCACAGAGCCTCTCACAACCTTTTGTGCCCCCAGGTTCT 114048
QY 541 TGGGAGTGAATGAATGATGATCATCATGGAATCATTTGGGGTCTCTAA 591
Db 114049 TGGGAGTGAATGAATGATGATCATCATGGAATCATTTGGGGTCTCTAA 114099

RESULT 15
CQ463286 CQ463286 506 bp DNA linear PAT 30-JAN-2004
LOCUS Sequence 6064 from Patent WO0192581.
DEFINITION
ACCESSION CQ463286

VERSION CQ463286.1 GI:41428905
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
AUTHORS Algate,P.A., Harlocker,S.L. and Jones,R.
TITLE Compositions and methods for the therapy and diagnosis of
JOURNAL ovariancancer
Patent: WO 0192581-A 6064 06-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source location/Qualifiers
1. 506
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 80.0%; Score 473; DB 6; length 506;
Best Local Similarity 98.6%; Pred. No. 5.3e-89;
Matches 488; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
QY 60 GTGAGTCCAGCAGGCTTCAAGAGCTACCAAGAGCCCTGCTCCGATCTCCCTAGACAA 119
Db 14 GCGGAGTCCAGCAGGCTTCAAGAGCTACCAAGAGCCCTGCTCCGATCTCCCTAGACAA 73
QY 120 AGTCCAGCGCAGCCTGGGCCCCGAGCAACCCAGCCTCCGAGCATGTCTCATCCATAA 179
Db 74 AGTCCAGCGC--CCTGGGCCCCGAGCAACCCAGCCTCCGAGCATGTCTCATCCATAA 131
QY 180 CACCTTCCAACAGCTGAGGCTGCACTTGCCCTGGCTCCGCCCCCTGCCCCCGGA 239
Db 132 CACCTTCCAACAGCTGAGGATGCACTTGCCCTGGCTCCGCCCCCTGCCCCCGGA 191
QY 240 GCCCTCTTCTGGGCGAGAGGATTCTCCCTGTCAAGCCACCATTTGGCTTATCTCAG 299
Db 192 GCCCTCTTCTGGGCGAGAGGATTCTCCCTGTCAAGCCACCATTTGGCTTATCTCAG 251
QY 300 GAGCTGACACCTCCATGATGAGGAGCTGAGCCCCCTCAGAATCCAGTACTCCCTTGG 359
Db 252 GAGCTGACACCTCCATGATGAGGAGCTGAGCCCCCTCAGAATCCAGTACTCCCTTGG 311
QY 360 CTCCAGAATGAAGTGCACCCCGAGCCTGATCCAGTCTTTTGAAGCTTGAGCTCCCG 419
Db 312 CTCCAGAATGAAGTGCACCCCGAGCCTGATCCAGTCTTTTGAAGCTTGAGCTCCCG 371
QY 420 GTAATTGGGGGACTCTGGCCTGATGACTTTCTTCTGACATTGACACATCTGCGGTAGA 479
Db 372 GTAATTGGGGGACTCTGGCCTGATGACTTTCTTCTGACATTGACACATCTGCGGTAGA 431
QY 480 AAGAGCCTGCACGGGCCCCACAGAGCCTCTCACAACCTCTTGTGCCCCCAGGTTT 539
Db 432 AAGAGCCTGCACGGGCCCCACAGAGCCTCTCACAACCTCTTGTGCCCCCAGGTTT 491
QY 540 TTGGAGTGAATGA 554
Db 492 TTGGAGTGAATGA 506

Search completed: February 6, 2006, 11:34:39
Job time : 3211 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 10:26:36 ; Search time 476 Seconds
(without alignments)
8274.860 Million cell updates/sec

Title: US-10-069-386A-1
Perfect score: 591
Sequence: 1 atggagggagcttgaagag.....aaatcattctggggtcctaa 591

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn19808:*
2: geneseqn19908:*
3: geneseqn20008:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	591	100.0	591 4	AAF28052	Aaf28052 Replicati
2	589.4	99.7	1325 4	AAK94884	Aak94884 Human ful
3	589.4	99.7	1325 12	ADL32051	Adl32051 Full leng
4	589.4	99.7	1515 13	ADQ85787	Adq85787 Human tum
5	589.4	99.7	1542 10	ADG33127	Adg33127 Human DNA
6	589.4	99.7	1542 14	ADZ70575	Adz70575 Human CDN
7	588.4	99.6	825 4	AAK91661	Aak91661 Human CDN
8	588.4	99.6	825 4	AAK93617	Aak93617 Human CDN
9	588.4	99.6	825 12	ADL30044	Adl30044 3' end of
10	588.4	99.6	825 12	ADL28088	Adl28088 5' end of
11	578.4	97.9	598 12	ACH91703	Ach91703 Human gen
12	576.2	97.5	1596 4	AAH34147	Aah34147 Human col
13	473	80.0	506 6	ABH83086	Abh83086 Human ova
14	186.6	31.6	205 6	ABH83261	Abh83261 Human ova
15	166	28.1	526 12	ACH77983	Ach77983 Human gen
16	47.4	8.0	2082 6	AAH41497	Aah41497 Mouse hae
c 17	46.4	7.9	51552 6	AAH96607	Aas96607 DNA encod
18	42.6	7.2	1603 8	ACC51095	Acc51095 Human Sho
19	42.4	7.2	511 12	ACH91786	Ach91786 Human gen

c	20	42.4	7.2	600	6	ABQ52497	Abq52497 Oligonuc1
c	21	42.4	7.2	600	6	ABQ52496	Abq52496 Oligonuc1
c	22	42.4	7.2	1850	8	ADA70325	Ada70325 Rice gene
c	23	41.8	7.1	32329	12	ADO51695	Ado51695 Streptomy
c	24	41.4	7.0	825	12	ADO00165	Ado00165 Novel hum
c	25	41.4	7.0	825	12	ADN98596	Adn98596 Novel hum
c	26	41	6.9	1925	2	AAH90924	Aah90924 Epatrein B
c	27	41	6.9	1926	3	AAA50254	Aaa50254 Epatrein B
c	28	41	6.9	1926	4	AAF82902	Aaf82902 BBV tethe
c	29	41	6.9	1926	10	ADK65580	Adk65580 Human her
c	30	41	6.9	1926	14	ADV68155	Adv68155 Epatrein-B
c	31	41	6.9	2046	14	AEH68224	Aeh68224 Epatrein-B
c	32	41	6.9	2580	3	AAA75454	Aaa75454 Nucleotid
c	33	41	6.9	2580	6	AAI64275	Aai64275 Epatrein-B
c	34	41	6.9	5452	2	AAH90923	Aah90923 Ant1-sens
c	35	41	6.9	8705	2	AAZ23778	Aaz23778 Vector pS
c	36	41	6.9	8705	12	ADM10659	Adm10659 Expressio
c	37	41	6.9	9482	12	ADP64415	Adp64415 Vector pC
c	38	41	6.9	9482	14	ADM05161	Adm05161 S/MAR vec
c	39	41	6.9	9600	2	AAV21683	Aav21683 Vector pI
c	40	41	6.9	10285	6	ABS71027	Abs71027 PCRp-Xa-F
c	41	41	6.9	10285	6	ABS66453	Abs66453 Plasmid p
c	42	41	6.9	10330	12	ADL67154	Adl67154 Plasmid p
c	43	41	6.9	10380	2	AAZ22248	Aaz22248 Nucleotid
c	44	41	6.9	10477	12	ADL67152	Adl67152 Plasmid p
c	45	41	6.9	10516	12	ADL67150	Adl67150 Plasmid p

ALIGNMENTS

RESULT 1	AAF28052	AAF28052 standard; DNA; 591 BP.
XX	XX	AAF28052;
AC	AC	23-MAY-2001 (first entry)
DT	DT	Replication protein A binding transcriptional activator 1 RBT1 gene.
XX	XX	RBT1; replication protein A binding transcriptional activator 1; RPA32;
KW	KW	gene therapy; apoptosis; cancer; leukaemia; ds.
XX	XX	Unidentified.
OS	OS	
XX	XX	
FH	FH	Key
FT	FT	CDS
FT	FT	Location/Qualifiers
FT	FT	1..591
FT	FT	/*tag= a
XX	XX	/product= "RBT1"
PN	PN	WO200114546-A2.
XX	XX	
PD	PD	01-MAR-2001.
XX	XX	
PF	PF	17-AUG-2000; 2000WO-CA000948.
XX	XX	
PR	PR	19-AUG-1999; 99US-0149472P.
XX	XX	
PA	PA	(TRAN-) CENT TRANSLATIONAL RES IN CANCER.
XX	XX	
PI	PI	Alaoui-Jamali MA, Cho JM;
XX	XX	
DR	DR	WPI; 2001-218447/22.
DR	DR	P-PSDB; AAB35402.
XX	XX	
PT	PT	Novel replication protein A binding transcriptional activator 1 gene,
PT	PT	useful for treating neoplastic disorders such as cancer and in gene
XX	XX	therapy.
XX	XX	Claim 1; Fig 1; 16bp; English.
PS	PS	
XX	XX	The present invention provides the protein and coding sequences of the
CC	CC	

Db 633 TGGGAGTGAATGAAC TGATCATCATGGAATCATTTGGGGTCCTAA 683

RESULT 3
ADL32051
ID ADL32051 standard; cDNA; 1325 BP.
XX
AC ADL32051;
XX
DT 20-MAY-2004 (first entry)
XX
DE Full length human cDNA clone SegID 4084.
XX
KW human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method; ss; gene.
XX
OS Homo sapiens.
XX
PN EPI396543-A2.
XX
PD 10-MAR-2004.
XX
PF 07-JUL-2000; 2003EP-00025638.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2004-204755/20.
XX
DR P-PSDB; ADL32052.
XX
PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
XX
PS Example 1; SEQ ID NO 4084; 1340pp; English.
XX
XX This invention relates to a novel primers useful for synthesising full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is a
CC full length human cDNA clone of the invention.
XX
SQ Sequence 1325 BP; 263 A; 371 C; 359 G; 332 T; 0 U; 0 Other;

Query Match 99.7%; Score 589.4; DB 12; Length 1325;
Best Local Similarity 99.8%; Pred. No. 1.4e-134;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGAGGAGGCTTGAAGAGAAACACTCTGATTGTGAAGAGAGAGAGAGGTGGAG 60
Db 93 ATGGTGGAGGCTTGAAGAGAAACACTCTGATTGTGAAGAGAGAGAGAGGTGGAG 152

QY 61 TGGAGTCCAGCAGGCTTCAAGCTACCAGCAGCCCTGCTCCGCATCTCCCTAGACAA 120
Db 153 TGGAGTCCAGCAGGCTTCAAGCTACCAGCAGCCCTGCTCCGCATCTCCCTAGACAA 212

QY 121 GTCCAGCGCAGCCTGGGGCCCCGAGCACCAGCCTCCGAGCGATGTCCTCATCCATAC 180
Db 213 GTCCAGCGCAGCCTGGGGCCCCGAGCACCAGCCTCCGAGCGATGTCCTCATCCATAC 272

QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTGCCCTGGCTCCCGCCCTGCCCCCGCAG 240
Db 273 ACCCTCCAACAGCTGCAGGCTGCACTTGCCCTGGCTCCCGCCCTGCCCCCGCAG 332

QY 241 CCCCTCTTCTGGGCGAGAGGATTCTCCCTGTCAAGCCACCATTTGGCTCTATCTCAGG 300
Db 333 CCCCTCTTCTGGGCGAGAGGATTCTCCCTGTCAAGCCACCATTTGGCTCTATCTCAGG 392

QY 301 GAGCTGGAACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTGAAGTCCCTTGGC 360
Db 393 GAGCTGGAACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTGAAGTCCCTTGGC 452

QY 361 CTCAGAAATGAAGTGCCACCCCAAGCTGATCCAGTCTTCTTAGAAGCTGTAGCTCCCGG 420
Db 453 CTCAGAAATGAAGTGCCACCCCAAGCTGATCCAGTCTTCTTAGAAGCTGTAGCTCCCGG 512

QY 421 TACTTGGGGGACTCTGGCCTTGATGACTTCTTCTTGACATTGACACATCTGCGGTAGAA 480
Db 513 TACTTGGGGGACTCTGGCCTTGATGACTTCTTCTTGACATTGACACATCTGCGGTAGAA 572

QY 481 AAGGAGCTGCACGGGCCCAAGAGCCTCTCACAACCTCTTGTGCCCCAGGTTCT 540
Db 573 AAGGAGCTGCACGGGCCCAAGAGCCTCTCACAACCTCTTGTGCCCCAGGTTCT 632

QY 541 TGGGAGTGAATGAATGATCATCATGGAATCATTTGGGGTCCTAA 591
Db 633 TGGGAGTGAATGAATGATCATCATGGAATCATTTGGGGTCCTAA 683

RESULT 4
ADQ85787
ID ADQ85787 standard; cDNA; 1515 BP.
XX
AC ADQ85787;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #2601.
XX
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.
XX
PA (GBTH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
PI Wu TD, Zhou Y;
XX
DR WPI; 2004-534300/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX
PS Claim 1; SEQ ID NO 2601; 5504pp; English.
XX

The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that has 80% sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-(c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence

CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.

XX Sequence 1515 BP; 328 A; 404 C; 385 G; 398 T; 0 U; 0 Other;

Query Match 99.7%; Score 589.4; DB 13; Length 1515;
Best Local Similarity 99.8%; Pred. No. 1.5e-134;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGGAGGAGCTTGAAGAGAAACACTGTATTTGAAGAGAGAGAGAGGTGGAG 60
DB 279 ATGAGTGGAGGCTTGAAGAGAAACACTGTATTTGAAGAGAGAGAGAGGTGGAG 338
QY 61 TGGAGTCCAGCAGGCGCTTCAAGACTAACGACGAGCCCTGCTCCGATCTCCCTAGACAA 120
DB 339 TGGAGTCCAGCAGGCGCTTCAAGACTAACGACGAGCCCTGCTCCGATCTCCCTAGACAA 398
QY 121 GTCCAGCGCAGCCTGGGGCCCCGAGCACCAGCCTCCGAGGATGCTCTCATCATTAAC 180
DB 399 GTCCAGCGCAGCCTGGGGCCCCGAGCACCAGCCTCCGAGGATGCTCTCATCATTAAC 458
QY 181 ACCCTCCACAGCTGACGCTGCACTTCCGCTGGCTCCGCCCCCTGCCCCCGAG 240
DB 459 ACCCTCCACAGCTGACGCTGCACTTCCGCTGGCTCCGCCCCCTGCCCCCGAG 518
QY 241 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTCAAGCACCATTGGCTCTATCTCAGG 300
DB 519 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTCAAGCACCATTGGCTCTATCTCAGG 578
QY 301 GAGCTGACACCTCCATGATGGAAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
DB 579 GAGCTGACACCTCCATGATGGAAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 638
QY 361 CTCAGAAATGAAGTGCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTTGAGCTCCGG 420
DB 639 CTCAGAAATGAAGTGCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTTGAGCTCCGG 698
QY 421 TACTTGGGGAGCTGTGCTGTGATGACTTCTTTCTGACATTGACATCTGCGTAGAA 480
DB 699 TACTTGGGGAGCTGTGCTGTGATGACTTCTTTCTGACATTGACATCTGCGTAGAA 758
QY 481 AAGGAGCTGCACGCGCCCAAGAGCCTCTCAAACTCTTCTGTGCCCAAGTTCT 540
DB 759 AAGGAGCTGCACGCGCCCAAGAGCCTCTCAAACTCTTCTGTGCCCAAGTTCT 818
QY 541 TGGGAGTGAATGAATGATGATCATCATGGAATCATTTGGGGTCTTAA 591
DB 819 TGGGAGTGAATGAATGATGATCATCATGGAATCATTTGGGGTCTTAA 869

RESULT 5
ADG33127
ID ADG33127 standard; DNA; 1542 BP.

AC ADG33127;
DT 26-FEB-2004 (first entry)
XX
DE Human DNA differentially expressed in patients with SLR SeqID451.

XX human; ds; autoimmune; chronic inflammatory disease; SLR;
KW systemic lupus erythematosus; rheumatoid arthritis; cholecystitis;
KW Sjogren's disease; CRIST syndrome; scleroderma; ankylosing spondylitis;
KW ulcerative colitis; primary sclerosing cholangitis; appendicitis;
KW diverticulitis; primary biliary sclerosis.

XX Homo sapiens.

OS WO2003090694-A2.

PN 06-NOV-2003.

PD 24-APR-2003; 2003WO-US013015.

PF 24-APR-2002; 2002US-00131827.

PR (EXPR-) EXPRESSION DIAGNOSTICS INC.

XX Wohlgemuth J, Fry K, Woodward R, Ly N;

XX WPI; 2003-877243/81.

XX Diagnosing or monitoring autoimmune and chronic inflammatory diseases,
PT such as rheumatoid arthritis, systemic lupus erythematosus, ulcerative
PT colitis, psoriasis and asthma by detecting the expression level of one or
PT more genes.

PS Claim 18; SEQ ID NO 451; 877bp; English.

XX This invention relates to novel methods for diagnosing and monitoring
CC autoimmune and chronic inflammatory diseases. Specifically, it refers to
CC the identification of genes that have a clinical utility as diagnostic
CC tools for the management of, in particular, patients with systemic lupus
CC erythematosus (SLE) or rheumatoid arthritis (RA). Accordingly, the
CC present invention describes a method for determining the levels of
CC multiple differentially expressed genes of a patient, in a concerted
CC manner, in order to achieve an improved diagnostic assay with sensitivity
CC and specificity for the disease in question. As such, these genes are
CC useful for the diagnosis of various other inflammatory disorders
CC including cholecystitis, Sjogren's disease, CRIST syndrome, scleroderma,
CC ankylosing spondylitis, ulcerative colitis, primary sclerosing
CC cholangitis, appendicitis, diverticulitis, and primary biliary sclerosis.
CC This polynucleotide is a DNA sequence representing human mRNA that is
CC differentially expressed in patients with SLE, used in an exemplification
CC of the invention.

XX Sequence 1542 BP; 346 A; 405 C; 388 G; 403 T; 0 U; 0 Other;

Query Match 99.7%; Score 589.4; DB 10; Length 1542;
Best Local Similarity 99.8%; Pred. No. 1.5e-134;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGGAGGAGCTTGAAGAGAAACACTGTATTTGAAGAGAGAGAGAGGTGGAG 60
DB 292 ATGAGTGGAGGCTTGAAGAGAAACACTGTATTTGAAGAGAGAGAGAGGTGGAG 351
QY 61 TGGAGTCCAGCAGGCGCTTCAAGACTAACGACGAGCCCTGCTCCGATCTCCCTAGACAA 120
DB 352 TGGAGTCCAGCAGGCGCTTCAAGACTAACGACGAGCCCTGCTCCGATCTCCCTAGACAA 411
QY 121 GTCCAGCGCAGCCTGGGGCCCCCAAGCACCAGCCTCCGAGGATGTCTCATCATTAAC 180

Db 412 GTCCAGCGCAGCCTGGGCCCCGAGCACCCAGCCTCCGCGCATGTCTCATCCATAC 471
Qy 181 ACCCTCCAACAGCTGCAGGCTGCACTTGCCTGCTCCGCCCTTGCCCTGCCCCCGAG 240
Db 472 ACCCTCCAACAGCTGCAGGCTGCACTTGCCTGCTCCGCCCTTGCCCTGCCCCCGAG 531
Qy 241 CCCCTCTTCTGGGCGAGAGATTTCCTCTGACGCCACCATTTGGCTCTATCTCAGG 300
Db 532 CCCCTCTTCTGGGCGAGAGATTTCCTCTGACGCCACCATTTGGCTCTATCTCAGG 591
Qy 301 GAGCTGACACCTCCATGATGGAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 592 GAGCTGACACCTCCATGATGGAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 651
Qy 361 CTCGAGAATGAAGTGCCACCCCGCTGATCCAGTCTTTAGAAAGCTCTGAGTCCCGG 420
Db 652 CTCGAGAATGAAGTGCCACCCCGCTGATCCAGTCTTTAGAAAGCTCTGAGTCCCGG 711
Qy 421 TACTTGGGGGACTCTGCTGATGACTTCTTTGACATTGACACATCTGCGTAGAA 480
Db 712 TACTTGGGGGACTCTGCTGATGACTTCTTTGACATTGACACATCTGCGTAGAA 771
Qy 481 AAGGAGCCTGCACGCGGCCCAACAGCCTCTCACAACTTTCTGTGCCCAAGTTCT 540
Db 772 AAGGAGCCTGCACGCGGCCCAACAGCCTCTCACAACTTTCTGTGCCCAAGTTCT 831
Qy 541 TGGGAGTGAATGAATGATGATCATCATGAAATCATTTGGGGTCTTAA 591
Db 832 TGGGAGTGAATGAATGATGATCATCATGAAATCATTTGGGGTCTTAA 882

RESULT 6
ADZ70575
ID ADZ70575 standard; cDNA; 1542 BP.
XX
AC ADZ70575;
XX
DT 30-JUN-2005 (first entry)
XX
DE Human cDNA from lung cancer marker gene RBT1.
XX
KW Tumor marker; ss; gene; lung tumor; cytostatic; neoplasm; expression;
KW DNA microarray.
XX
OS Homo sapiens.
XX
PN WO2005032495-A2.
XX
PD 14-APR-2005.
XX
PF 01-OCT-2004; 2004WO-US034163.
XX
PR 03-OCT-2003; 2003US-0508355P.
XX
PA (FARB) BAYER PHARM CORP.
XX
PI Taylor I, Pauloski NR, Bigwood D;
XX
DR WPI; 2005-285325/29.
DR P-PSDB; ADZ70576.
XX
XX Providing a patient diagnosis for lung cancer comprises comparing the
PT level of expression of genes or gene products in a biological sample from
PT the patient with that from a normal individual.
XX
XX Claim 2; SEQ ID NO 260; 60pp; English.
XX
CC The invention relates to providing a patient diagnosis for lung cancer
CC comprising comparing the level of expression of genes or gene products in
CC a biological sample from the patient with the level of expression of
CC genes or gene products in a biological sample from a normal individual.
CC Also included are distinguishing between normal and disease tissues,

CC monitoring the response of a patient being treated for lung cancer by
CC administering an anti-cancer agent, identifying a compound useful for the
CC treatment of lung cancer and an array for distinguishing between normal
CC and disease tissues (comprising 2 or more probes corresponding to 2 or
CC more genes selected from any of the 200 nucleotide sequences given in the
CC specification, or 2 or more polypeptides comprising any of the 200 amino
CC acid sequences given in the specification). In providing a patient
CC diagnosis for lung cancer, one or more genes are selected from any of the
CC 200 nucleotide sequences as mentioned in the specification, or one or
CC more gene products are polypeptides selected from any of the 20 amino
CC acid sequences mentioned in the specification. The methods are useful for
CC detecting and treating lung cancer. These may also be used for designing,
CC identifying and optimizing therapeutics for cancer. The present sequence
CC represents a cDNA from one of the 200 lung cancer marker genes. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 1542 BP; 346 A; 405 C; 388 G; 403 T; 0 U; 0 Other;

Query Match 99.7%; Score 589.4; DB 14; Length 1542;

Best Local Similarity 99.8%; Pred. No.1.5e-134;

Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAGGAGGAGGCTTGAAGAGAAACACTGTGATTGGAAGAGAGAGAGAGTGGAG 60
Db 292 ATGATGGAGAGGCTTGAAGAGAAACACTGTGATTGGAAGAGAGAGAGAGTGGAG 351
Qy 61 TGGAGTCCAGCAGGCTTCAAGCTAACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 120
Db 352 TGGAGTCCAGCAGGCTTCAAGCTAACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 411
Qy 121 GTCCAGCGCAGCCTGGGCCCCCGAGACCCAGCCTCCGCGCATGTCTCATCCATAC 180
Db 412 GTCCAGCGCAGCCTGGGCCCCCGAGACCCAGCCTCCGCGCATGTCTCATCCATAC 471
Qy 181 ACCCTCCAACAGCTGCAGGCTGCACTTGCTGCTGCTCCGCCCTGCCCCCGAG 240
Db 472 ACCCTCCAACAGCTGCAGGCTGCACTTGCTGCTGCTCCGCCCTGCCCCCGAG 531
Qy 241 CCCCTCTTCTGGGCGAGAGATTTCCTCTGACGCCACCATTTGGCTCTATCTCAGG 300
Db 532 CCCCTCTTCTGGGCGAGAGATTTCCTCTGACGCCACCATTTGGCTCTATCTCAGG 591
Qy 301 GAGCTGACACCTCCATGATGGAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 592 GAGCTGACACCTCCATGATGGAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 651
Qy 361 CTCGAGAATGAAGTGCCACCCCGCTGATCCAGTCTTTAGAAAGCTCTGAGTCCCGG 420
Db 652 CTCGAGAATGAAGTGCCACCCCGCTGATCCAGTCTTTAGAAAGCTCTGAGTCCCGG 711
Qy 421 TACTTGGGGGACTCTGCTGATGACTTCTTTGACATTGACACATCTGCGTAGAA 480
Db 712 TACTTGGGGGACTCTGCTGATGACTTCTTTGACATTGACACATCTGCGTAGAA 771
Qy 481 AAGGAGCCTGCACGCGGCCCAACAGAGCCTCTCACAACTTTCTGTGCCCCAGGTTCT 540
Db 772 AAGGAGCCTGCACGCGGCCCAACAGAGCCTCTCTCACAACTTTCTGTGCCCCAGGTTCT 831
Qy 541 TGGGAGTGAATGAATGATGATCATCATGAAATCATTTGGGGTCTTAA 591
Db 832 TGGGAGTGAATGAATGATGATCATCATGAAATCATTTGGGGTCTTAA 882

RESULT 7
AAK91661
ID AAK91661 standard; cDNA; 825 BP.
XX
XX AAK91661;
AC
XX
DT 06-NOV-2001 (first entry)
XX

DE Human cDNA 5'-end sequence, SEQ ID NO: 121.
XX
XX Human, full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000BP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 2; SEQ ID NO 121; 1380bp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is the nucleotide sequence of the 5'-end of
CC a cDNA provided in the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in CD-
CC ROM format directly from EPO
XX
SQ Sequence 825 BP; 160 A; 247 C; 230 G; 183 T; 0 U; 5 Other;

Query Match 99.6%; Score 588.4; DB 4; Length 825;
Best Local Similarity 99.7%; Pred. No. 2.2e-134;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGGAGGCTTGAAGAGAAACACTGTGATTGGAGAGAGAGAGAGGTTGGAG 60
DB 93 ATGTGGAGGCTTGAAGAGAAACACTGTGATTGGAGAGAGAGAGAGGTTGGAG 152

QY 61 TGGAGTCCAGCAGGCGCTTTCAGAGCTACCAAGAGCCCTGCTCCGATCTCCCTAGACAA 120
DB 153 TGGAGTCCAGCAGGCGCTTTCAGAGCTACCAAGAGCCCTGCTCCGATCTCCCTAGACAA 212

QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGAGGATGTCTCATCCATAAC 180
DB 213 GTCCAGCGCAGCCTGGGCCCCCGAGCACCAGCCTCCGAGGATGTCTCATCCATAAC 272

QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTGGCTGGCTCCCGCCCTGCCCCCGAG 240
DB 273 ACCCTCCAACAGCTGCAGGCTGCACTTGGCTGGCTCCCGCCCTGCCCCCGAG 332

QY 241 CCCCTCTTCTGTTGGGCGAGGAGGATTTCCTCTGTACGCCACCATTTGGCTATCTCAAG 300
DB 333 CCCCTCTTCTGTTGGGCGAGGAGGATTTCCTCTGTACGCCACCATTTGGCTATCTCAAG 392

QY 301 GAGCTGACACCTCCATGATGGAAGTGAAGCCCTCAGAAATCAGTACTCCCTTGGC 360
DB 393 GAGCTGACACCTCCATGATGGAAGTGAAGCCCTCAGAAATCAGTACTCCCTTGGC 452

QY 361 CTCGAATGAAGTGCACCCAGCCTGATCCAGTCTTCTAGAGCTCTGAGTCCCGG 420

DB 453 CTCGAATGAAGTGCACCCAGCCTGATCCAGTCTTCTTAGAGCTTGAGCTCCCGG 512
QY 421 TACTTGGGAGCTTGCCCTGGATGACTTCTTCTGACATTTGACACATCTGGCGTAGAA 480
DB 513 TACTTGGGAGCTTGCCCTGGATGACTTCTTCTGACATTTGACACATCTGGCGTAGAA 572
QY 481 AAGAGCCTGCACGGGCCCCCAGAGAGCCTCCTCACAACCTCTTGTGCCCAAGTTCT 540
DB 573 AAGAGCCTGCACGGGCCCCCAGAGAGCCTCCTCACAACCTCTTGTGCCCAAGTTCT 632
QY 541 TGGAGTGAATGAATGGATCATCATGTGAATCATTTCTGGGGTCTAA 591
DB 633 TGGAGTGAATGAATGGATCATCATGTGAATCATTTCTGGGGTCTAA 683

RESULT 8
AAK93617
ID AAK93617 standard; cDNA; 825 BP.
XX
XX AAK93617;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human cDNA clone representative sequence, SEQ ID NO: 2077.
XX
XX Human, full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
XX Homo sapiens.
XX
XX EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000BP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Example 11; SEQ ID NO 2077; 1380bp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence was used as the representative sequence
CC from a human clone which was used in homology searches to identify the
CC clone. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in CD-ROM format directly from
CC EPO
XX
SQ Sequence 825 BP; 160 A; 247 C; 230 G; 183 T; 0 U; 5 Other;

Query Match 99.6%; Score 588.4; DB 4; Length 825;
Best Local Similarity 99.7%; Pred. No. 2.2e-134;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAGGAGGCTTGAAGAGAAACACTGTGATTGGAGAGAGAGAGGTTGGAG 60

Db 93 ATGGTGGAGGCTTGAAGAGAAACACTCTGATTGTGAAGAGAGAGAGAGAGTGGAG 152
QY 61 TGGAGTCCAGCAGGCGCTTCAAGAGTACCAGCAAGCCCTGCTCCGATCTCCCTAGACAA 120
Db 153 TGGAGTCCAGCAGGCGCTTCAAGAGTACCAGCAAGCCCTGCTCCGATCTCCCTAGACAA 212
QY 121 GTCCAGCGCAGCCTGGGCCCCGAGCACCAGCCTCCGAGGATGTCTCATCCATAAC 180
Db 213 GTCCAGCGCAGCCTGGGCCCCGAGCACCAGCCTCCGAGGATGTCTCATCCATAAC 272
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTCCCTGGCTCCGCCCCCTGCCCCCGAG 240
Db 273 ACCCTCCAACAGCTGCAGGCTGCACTTCCCTGGCTCCGCCCCCTGCCCCCGAG 332
QY 241 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTCAAGCCAGATTGGCTTATCCTCAG 300
Db 333 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTCAAGCCAGATTGGCTTATCCTCAG 392
QY 301 GAGCTGGACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 393 GAGCTGGACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 452
QY 361 CTCAGAAATGAAGTCCACCCCAAGCTGTATCCAGTCTTGAAGCTCTGAGCTCCCG 420
Db 453 CTCAGAAATGAAGTCCACCCCAAGCTGTATCCAGTCTTGAAGCTCTGAGCTCCCG 512
QY 421 TACTTGGGGGACTGTGCTGTGAATGACTTCTTCTGACATGACACATCTGCGGTAGA 480
Db 513 TACTTGGGGGACTGTGCTGTGAATGACTTCTTCTGACATGACACATCTGCGGTAGA 572
QY 481 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTTCTGTGCCCAAGTTCT 540
Db 573 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTTCTGTGCCCAAGTTCT 632
QY 541 TGGGAGTGAATGAATGATGATCATCATGGAATCATTTGGGGTCTTAA 591
Db 633 TGGGAGTGAATGAATGATGATCATCATGGAATCATTTGGGGTCTTAA 683

RESULT 9
ADL30044
ID ADL30044 standard; cDNA; 825 BP.
XX
AC ADL30044;
XX
DT 20-MAY-2004 (first entry)
XX
DE 3' end of a representative human cDNA cluster SeqID 2077.
XX
KM human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method; ss.
XX
OS Homo sapiens.
XX
PN BP1396543-A2.
XX
PD 10-MAR-2004.
XX
PF 07-JUL-2000; 2003BP-00025638.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000BP-00114089.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2004-204755/20.
XX
PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full

PT length human cDNAs.
XX
XX Example 18; SEQ ID NO 2077; 1340bp; English.
XX
CC This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is the
CC 3' end of a representative human DNA cluster of the invention.
XX
SQ Sequence 825 BP; 160 A; 247 C; 230 G; 183 T; 0 U; 5 Other;
Query Match 99.6%; Score 588.4; DB 12; Length 825;
Best Local Similarity 99.7%; Pred. No. 2.2e-134;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGGGAAGGCTTGAAGAGAAACACTCTGATTGTGAAGAGAGAGAGAGTGGAG 60
Db 93 ATGGTGGAGGCTTGAAGAGAAACACTCTGATTGTGAAGAGAGAGAGAGTGGAG 152
QY 61 TGGAGTCCAGCAGGCGCTTCAAGAGTACCAGCAAGCCCTGCTCCGATCTCCCTAGACAA 120
Db 153 TGGAGTCCAGCAGGCGCTTCAAGAGTACCAGCAAGCCCTGCTCCGATCTCCCTAGACAA 212
QY 121 GTCCAGCGCAGCCTGGGCCCCGAGCACCAGCCTCCGAGGATGTCTCATCCATAAC 180
Db 213 GTCCAGCGCAGCCTGGGCCCCGAGCACCAGCCTCCGAGGATGTCTCATCCATAAC 272
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTCCCTGGCTCCGCCCCCTGCCCCCGAG 240
Db 273 ACCCTCCAACAGCTGCAGGCTGCACTTCCCTGGCTCCGCCCCCTGCCCCCGAG 332
QY 241 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTCAAGCCAGATTGGCTTATCCTCAG 300
Db 333 CCCCTCTTCTGGGCGAGAGATTCTCCCTGTCAAGCCAGATTGGCTTATCCTCAG 392
QY 301 GAGCTGGACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 393 GAGCTGGACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 452
QY 361 CTCAGAAATGAAGTCCACCCCAAGCTGTATCCAGTCTTGAAGCTCTGAGCTCCCG 420
Db 453 CTCAGAAATGAAGTCCACCCCAAGCTGTATCCAGTCTTGAAGCTCTGAGCTCCCG 512
QY 421 TACTTGGGGGACTGTGCTGTGAATGACTTCTTCTGACATGACACATCTGCGGTAGA 480
Db 513 TACTTGGGGGACTGTGCTGTGAATGACTTCTTCTGACATGACACATCTGCGGTAGA 572
QY 481 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTTCTGTGCCCAAGTTCT 540
Db 573 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTTCTGTGCCCAAGTTCT 632
QY 541 TGGGAGTGAATGAATGATGATCATCATGGAATCATTTGGGGTCTTAA 591
Db 633 TGGGAGTGAATGAATGATGATCATCATGGAATCATTTGGGGTCTTAA 683

RESULT 10
ADL28088
ID ADL28088 standard; cDNA; 825 BP.
XX
AC ADL28088;
XX
DT 20-MAY-2004 (first entry)
XX
DE 5' end of a human cDNA molecule SeqID 121.
XX
KM human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method; ss.

XX Homo sapiens.
OS
XX
PN BP1396543-A2.
XX
PD 10-MAR-2004.
XX
PF 07-JUL-2000; 2003BP-00025638.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2004-204755/20.
XX
PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
XX
PS Disclosure; SEQ ID NO 121; 1340bp; English.
XX
CC This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is the
CC 5' end of a full length human cDNA sequence of the invention.
XX
SQ Sequence 825 BP; 160 A; 247 C; 230 G; 183 T; 0 U; 5 Other;

Query Match 99.6%; Score 588.4; DB 12; Length 825;
Best Local Similarity 99.7%; Pred. No. 2.2e-134;
Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGGGAGGCTTGAAGAGGAAACACTGTATTGGAAAGAGAGAGAGAGGTGGAG 60
DB 93 ATGTGGAGGCTTGAAGAGGAAACACTGTATTGGAAAGAGAGAGAGAGGTGGAG 152

QY 61 TGGAGTCCAGCAGGCGCTTCAAGCTACAGCAAGCGCTGCTCCGCATCTCCCTAGCAAA 120
DB 153 TGGAGTCCAGCAGGCGCTTCAAGCTACAGCAAGCGCTGCTCCGCATCTCCCTAGCAAA 212

QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCAACCCAGCCTCCGAGGCATGTCTCATCCATAAC 180
DB 213 GTCCAGCGCAGCCTGGGCCCCCGAGCAACCCAGCCTCCGAGGCATGTCTCATCCATAAC 272

QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTCGCTGGCTCCGCCCCCTGCCCCCGAG 240
DB 273 ACCCTCCAACAGCTGCAGGCTGCACTTCGCTGGCTCCGCCCCCTGCCCCCGAG 332

QY 241 CCCCTCTTCTGGGCGAGGAGATTCTCCCTGTACGCCACATTGGCTTATCTCAGG 300
DB 333 CCCCTCTTCTGGGCGAGGAGATTCTCCCTGTACGCCACATTGGCTTATCTCAGG 392

QY 301 GAGCTGACACCTCCATGATGGAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
DB 393 GAGCTGACACCTCCATGATGGAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 452

QY 361 CTCAGAATGAAGTGCCACCCGAGCTGATCCAGTCTTTTGAAGACTCTGAGTCCCGG 420
DB 453 CTCAGAATGAAGTGCCACCCGAGCTGATCCAGTCTTTTGAAGACTCTGAGTCCCGG 512

QY 421 TACTTGGGGGACTCTGGCCTGATGACTTTTCTTGACATTGACACATCTGCGGTAGA 480
DB 513 TACTTGGGGGACTCTGGCCTGATGACTTTTCTTGACATTGACACATCTGCGGTAGA 572

QY 481 AAGAGCCTGCACGGGCCCCCACCAGAGCCTCTCTCAACAACCTCTTGTGCCCCAGGTCT 540
DB 573 AAGAGCCTGCACGGGCCCCCACCAGAGCCTCTCTCAACAACCTCTTGTGCCCCAGGTCT 632

QY 541 TGGAGTGAATGAATGATCATCATCATGAAATCATTTCTGGGTCCTAA 591
DB 633 TGGAGTGAATGAATGATCATCATCATGAAATCATTTCTGGGTCCTAA 683

RESULT 11
ACH91703
ID ACH91703 standard; DNA; 598 BP.
XX
AC ACH91703;
XX
DT 29-JUL-2004 (first entry)
XX
DB Human genome derived single exon probe #24898.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 1; SEQ ID NO 24898; 80bp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon probes is separately
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, and a computer-readable
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their

CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
XX

SQ Sequence 598 BP; 123 A; 197 C; 158 G; 120 T; 0 U; 0 Other;

Query Match 97.9%; Score 578.4; DB 12; Length 598;
Best Local Similarity 99.7%; Pred. No. 5.9e-132;
Matches 590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATGGAGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 60
Db 7 ATGCTGGAGAGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 66
QY 61 TGGAGTCCAGCAGCGCTTCAGAGCTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 120
Db 67 TGGAGTCCAGCAGCGCTTCAGAGCTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 126
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCAACCAGCCTCCGAGGATGTCTCATCCATAAC 180
Db 127 GTCCAGCGCAGCCTGGGCCCCCGAGCAACCAGCCTCCGAGGATGTCTCATCCATAAC 186
QY 181 ACCCTCCAACAGCTGAGGCTGCACTTGCTGGCTCCCGCCCTGCCCCG-CCCCCGA 239
Db 187 ACCCTCCAACAGCTGAGGCTGCACTTGCTGGCTCCCGCCCTGCCCCG-CCCCCGA 246
QY 240 GCGCCCTCTTCTTGGGCGAGAGATTCTCCCTGTACGCCACCATTTGGCTTATCTCAG 299
Db 247 GCGCCCTCTTCTTGGGCGAGAGATTCTCCCTGTACGCCACCATTTGGCTTATCTCAG 306
QY 300 GAGAGTGAACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTGACTCCCTTG 359
Db 307 GAGAGTGAACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTGACTCCCTTG 366
QY 360 CCTCCAGAATGAAGTCCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTGAGCTCCG 419
Db 367 CCTCCAGAATGAAGTCCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTGAGCTCCG 426
QY 420 GTACTTGGGGGACTCTGGCCTTGATGACTTCTTTTGACATGACATCTGCGGTAG 479
Db 427 GTACTTGGGGGACTCTGGCCTTGATGACTTCTTTTGACATGACATCTGCGGTAG 486
QY 480 AAAGAGCCTGACGCGGCCCAAGAGCCTCTCACAACCTTCTGTGCCCCAGGTT 539
Db 487 AAAGAGCCTGACGCGGCCCAAGAGCCTCTCTCACAACCTTCTGTGCCCCAGGTT 546
QY 540 TTGGGAGTGAATGAATGATCATCATGGAATCATTTCTGGGGTCTTA 591
Db 547 TTGGGAGTGAATGAATGATCATCATGGAATCATTTCTGGGGTCTTA 598

RESULT 12
AAH34147
ID AAH34147 standard; cDNA; 1596 BP.

XX AAH34147;
XX
DT 03-SEP-2001 (first entry)
XX
DB Human colon cancer antigen encoding cDNA SEQ ID NO:1229.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KM colorectal carcinoma; ss.
XX
OS Homo sapiens.
XX

PN WO200122920-A2.
XX
XX 05-APR-2001.
PD
XX
PF 28-SEP-2000; 2000WO-US026524.
XX
XX 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
XX WPI; 2001-235357/24.
DR
DR P-PSDB; AAG74742.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX
PS Claim 1; Page 3003-3004; 9803pp; English.

AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patient's own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX
SQ Sequence 1596 BP; 354 A; 418 C; 410 G; 409 T; 0 U; 5 Other;

Query Match 97.5%; Score 576.2; DB 4; Length 1596;
Best Local Similarity 99.3%; Pred. No. 2.6e-131;
Matches 587; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY 1 ATGGAGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 60
Db 345 ATGGAGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 404
QY 61 TGGAGTCCAGCAGGCTTCAGAGCTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 120
Db 405 TGGAGTCCAGCAGGCTTCAGAGCTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 464
QY 121 GTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCTCCGAGGATGTCTCATCCATAAC 180
Db 465 GTCCAGCGC--CCTGGGCCCCCGAGCACCCAGCCTCCGAGGATGTCTCATCCATAAC 522
QY 181 ACCCTCCAACAGCTGAGGCTGCACTTGCGCTGGCTCCGCCCTGCCCCCGAG 240
Db 523 ACCCTCCAACAGCTGAGGCTGCACTTGCGCTGGCTCCGCCCTGCCCCCGAG 582
QY 241 CCCCTCTTCTTGGGCGAGAGATTCTCCCTGTACGCCACCATTTGGCTTATCTCAGG 300
Db 583 CCCCTCTTCTTGGGCGAGAGATTCTCCCTGTACGCCACCATTTGGCTTATCTCAGG 642
QY 301 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTTGAATCCAGTGACTCCCTTGGC 360
Db 643 GAGCTGACACCTCCATGATGGGACTGAGCCCCCTTGAATCCAGTGACTCCCTTGGC 702
QY 361 CTCCAGAATGAAGTGCACCCCAAGCCTGATCCAGTCTTCTTGAAGCTCTGAGCTCCGG 420
Db 703 CTCCAGAATGAAGTGCACCCCAAGCCTGATCCAGTCTTCTTGAAGCTCTGAGCTCCGG 762

QY 421 TACTTGGGGACTCTGGCTGGATGACTTCTTTCTGACATTGACACATCTGCGGTAGAA 480
|||||
Db 763 TACTTGGGGACTCTGGCTGGATGACTTCTTTCTGACATTGACACATCTGCGGTAGAA 822
QY 481 AAGGAGCCTGCACGGGCCCCCAGAGAGCCTCTTCTGAGAGCTTCTGAGAGGTTCT 540
|||||
Db 823 AAGGAGCCTGCACGGGCCCCCAGAGAGCCTCTTCTGAGAGGTTCT 882
QY 541 TGGAGTGAATGAATGATCATCATGAAATCATCTGGGTCCTAA 591
|||||
Db 883 TGGAGTGAATGAATGATCATCATGAAATCATCTGGGTCCTAA 933

RESULT 13

ABL83086
ID ABL83086 standard; cDNA; 506 BP.
XX ABL83086;
AC
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:6064.
XX
XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
XX Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US017756.
PR
XX 26-MAY-2000; 2000US-0207484P.
XX
XX (CORI-) CORIXA CORP.
PA
XX
PI Algate PA, Harlocker SL, Jones R;
XX
XX WPI; 2002-122075/16.
DR
XX
XX
PT
PT
PT
PT
XX
XX
PS
XX
XX

Composition for therapy and diagnosis of ovarian cancer comprising polypeptide of a ovarian tumor polypeptide, polynucleotide encoding polypeptide, antibody specific to polypeptide or T cell expressing polypeptide.

PS Claim 1; SEQ ID NO 6064; 489pp; English.
XX
XX The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumor
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC or antigen presenting cells that express (II). (I) has cytostatic
CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC detecting ovarian cancer in a patient's biological sample preferably
CC serum or ovarian tissue. The method comprises contacting a biological
CC sample from a patient with (IV), detecting the amount of polynucleotide
CC hybridising to (IV) and comparing the amount to a predetermined cutoff
CC value and thereby detecting ovarian cancer in the patient, where the
CC amount of polynucleotide hybridising to (IV) is detected preferably by
CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC useful for stimulating and/or expanding T cells specific for an ovarian
CC tumour protein comprising contacting T cells with (III) or (II). (III) is
CC useful in design and preparation of ribozyme molecules for inhibiting
CC expression of the tumour polypeptides and proteins in tumour cells; and
CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC library using well known techniques
XX
XX Sequence 506 BP; 98 A; 184 C; 124 G; 100 T; 0 U; 0 Other;

Query Match 80.0%; Score 473; DB 6; Length 506;

Best Local Similarity 98.6%; Pred. No. 4.1e-106;
Matches 488; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 60 GTGAGTCCAGCAGGCTTACAGAGCTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 119
|||||
Db 14 GCGAGTCCAGCAGGCTTACAGAGCTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 73
QY 120 AGTCAGCGCAGCTGGAGCCCCCAGACCCAGCCTCCGAGGATGTCCTCATCATAA 179
|||||
Db 74 AGTCAGCGC--CCTGGAGCCCCCAGACCCAGCCTCCGAGGATGTCCTCATCATAA 131
QY 180 CACCCTCCAACAGCTGACAGCTGCACTTGCCTGCTCCGCCCTGCTGCCCTCCGGA 239
|||||
Db 132 CACCCTCCAACAGCTGACAGCTGCACTTGCCTGCTCCGCCCTGCTGCCCTCCGGA 191
QY 240 GCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTACGCCACCATTTGCTCTATCCTCAG 299
|||||
Db 192 GCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTACGCCACCATTTGCTCTATCCTCAG 251
QY 300 GAGAGCTGACACCTTCATGATGGAGCTGAGGCCCTCAGAATCCAGTACTCCCTTGG 359
|||||
Db 252 GAGAGCTGACACCTTCATGATGGAGCTGAGGCCCTCAGAATCCAGTACTCCCTTGG 311
QY 360 CCTCCAGATGAATGAGTCCACCCAGCCTGATCAGTCTTTTGAAGCTCTGAGCTCCCG 419
|||||
Db 312 CCTCCAGATGAATGAGTCCACCCAGCCTGATCAGTCTTTTGAAGCTCTGAGCTCCCG 371
QY 420 GTACTTGGGGAGCTTGGCCTGGATGACTTCTTTCTGACATTGACACATCTGCGGTAGA 479
|||||
Db 372 GTACTTGGGGAGCTTGGCCTGGATGACTTCTTTCTGACATTGACACATCTGCGGTAGA 431
QY 480 AAGGAGCCTGCACGGGCCCCCAGAGAGCCTCTTCTGAGAGGTTCT 539
|||||
Db 432 AAGGAGCCTGCACGGGCCCCCAGAGAGCCTCTTCTGAGAGGTTCT 491
QY 540 TTGGAGTGAATGA 554
|||||
Db 492 TTGGAGTGAATGA 506

RESULT 14

ABL83261
ID ABL83261 standard; cDNA; 205 BP.
XX
AC ABL83261;
XX
DT 17-MAY-2002 (first entry)
XX
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:6239.
XX
XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
XX Homo sapiens.
OS
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US017756.
PR
XX 26-MAY-2000; 2000US-0207484P.
XX
XX (CORI-) CORIXA CORP.
PA
XX
PI Algate PA, Harlocker SL, Jones R;
XX
XX WPI; 2002-122075/16.
DR
XX
XX
PT
PT
PT
PT
XX
XX

Composition for therapy and diagnosis of ovarian cancer comprising polypeptide of a ovarian tumor polypeptide, polynucleotide encoding polypeptide, antibody specific to polypeptide or T cell expressing polypeptide.

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 10:41:17 ; Search time 3756 Seconds
(without alignments)
7361.865 Million cell updates/sec

Title: US-10-069-386A-1
Perfect score: 591
Sequence: 1 atggaggaggagctgaagag.....aatcattctgggtcctaa 591

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	gb_est1:*
2:	gb_est2:*
3:	gb_est3:*
4:	gb_htc:*
5:	gb_est4:*
6:	gb_est5:*
7:	gb_est6:*
8:	gb_est7:*
9:	gb_gss1:*
10:	gb_gss2:*
11:	gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	589.4	99.7	591	10 AY408336	AY408336 Homo sapi
2	589.4	99.7	729	7 CN304585	CN304585 170005005
3	589.4	99.7	742	7 CN304584	CN304584 170005979
4	589.4	99.7	1038	3 BM560308	BM560308 AGENCOURT
5	589.4	99.7	1298	4 CR605181	CR605181 full-leng
6	589.4	99.7	1318	4 CR601245	CR601245 full-leng
7	582.6	98.6	1056	5 BU157342	BU157342 AGENCOURT
8	580.4	98.2	646	7 CN304590	CN304590 170006001
9	575.8	97.4	1017	5 BX398356	BX398356 BX398356
10	575.2	97.3	712	7 CN304594	CN304594 170005326
11	567.8	96.1	873	5 BU508982	BU508982 AGENCOURT
12	560.8	94.9	970	7 COS79535	COS79535 ILLUMIGEN
13	559.8	94.7	668	2 BG720189	BG720189 602692233
14	558	94.4	1001	5 BX362283	BX362283 BX362283
15	557	94.2	1045	3 BM546306	BM546306 AGENCOURT
16	529	89.5	612	5 BX113979	BX113979 BX113979
17	528.8	89.5	649	3 BM716519	BM716519 UI-E-EJO-
18	527	89.2	575	7 CV028541	CV028541 7081 Full
19	523.4	88.6	642	2 BG105954	BG105954 602290144
20	517.8	87.6	649	7 CN304591	CN304591 170004245
21	501	84.8	838	6 CB995481	CB995481 AGENCOURT
22	499.6	84.5	705	2 BG527376	BG527376 602557362

23	494.4	83.7	598	7 CN304593	CN304593 170005977
24	493.8	83.6	785	8 DR773605	DR773605 ILLUMIGEN
25	482.4	81.6	568	3 BP262533	BP262533 BP262533
26	479	81.0	784	5 BW967346	BW967346 BW967346
27	479	81.0	800	3 BP463192	BP463192 BP463192
28	479	81.0	838	8 DN121997	DN121997 1122166 M
29	479	81.0	858	3 BP433413	BP433413 BP433413
30	477	80.7	784	3 CN304589	CN304589 170005998
31	476.4	80.6	582	3 BP320100	BP320100 BP320100
32	476	80.5	695	8 DN127988	DN127988 1142035 M
33	473	80.0	506	1 AA482412	AA482412 zt34d02.r
34	471	79.7	671	7 CN792897	CN792897 4127835 B
35	471	79.7	679	7 CK837132	CK837132 4061918 B
36	471	79.7	704	7 CK967446	CK967446 4082464 B
37	471	79.7	718	7 CK832545	CK832545 4056138 B
38	471	79.7	730	8 DN272726	DN272726 LIB30321
39	469.4	79.4	598	3 BM848331	BM848331 K-EST0128
40	467.8	79.2	815	7 CR850218	CR850218 CR850218
41	467.4	79.1	473	10 AY408337	AY408337 Pan trogl
42	459	77.7	1196	3 BM805649	BM805649 AGENCOURT
43	446.4	75.5	662	7 CK969781	CK969781 4085423 B
44	435.6	73.7	663	8 DN276725	DN276725 1155912 M
45	435.4	73.7	526	8 DN996881	DN996881 TC111917

ALIGNMENTS

RESULT 1
AY408336
LOCUS
DEFINITION Homo sapiens HCM3190 gene, VIRTUAL TRANSCRIPT, partial sequence,
ACCESSION AY408336
VERSION AY408336.1 GI:39764307
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
PUBMED Science 302 (5652), 1960-1963 (2003)
REFERENCES
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE
Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
location/Qualifiers
source
1..591
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>591
/locus_tag="HCM3190"

ORIGIN
Query Match 99.7%; Score 589.4; DB 10; Length 591;
Best Local Similarity 99.8%; Pred. No. 2e-135;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAGAGGAGCTTGAAGAGAAACACTGTATTTGGAAGAGAGAGAGAGGTGGAG 60
    |||
Db 1 ATGTTGGAGAGCTTGAAGAGAAACACTGTATTTGGAAGAGAGAGAGAGGTGGAG 60
QY 61 TGGAGTCCAGCAGGCGCTTCAAGACTAACAGCAAGCCCTGCTCCGCATCTCCCTAGACAA 120
    |||
Db 61 TGGAGTCCAGCAGGCGCTTCAAGACTAACAGCAAGCCCTGCTCCGCATCTCCCTAGACAA 120
QY 121 GTCCAGCGCAGCCTGGGCCCCGAGACACCCAGCCTCCGAGCATGTCTCATCAATAAC 180
    |||
Db 121 GTCCAGCGCAGCCTGGGCCCCGAGACACCCAGCCTCCGAGCATGTCTCATCAATAAC 180
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTGCCCTGGCTCCGCCCTGCTGCCCTGCCCGAG 240
    |||
Db 181 ACCCTCCAACAGCTGCAGGCTGCACTTGCCCTGGCTCCGCCCTGCTGCCCGAG 240
QY 241 CCCCTCTTCTGCGGCGAGGAGATTCTCCCTGTCAAGCACCATTGGCTCTATCCTCAGG 300
    |||
Db 241 CCCCTCTTCTGCGGCGAGGAGATTCTCCCTGTCAAGCACCATTGGCTCTATCCTCAGG 300
QY 301 GAGCTGGAACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
    |||
Db 301 GAGCTGGAACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
QY 361 CTCGAGAATGAAGTGCCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTTGAGCTCCGG 420
    |||
Db 361 CTCGAGAATGAAGTGCCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTTGAGCTCCGG 420
QY 421 TACTTGGGGGACTCTGCGCTGATGACTTCTTTCTGGAACATGACACATCTGCGTAGAA 480
    |||
Db 421 TACTTGGGGGACTCTGCGCTGATGACTTCTTTCTGGAACATGACACATCTGCGTAGAA 480
QY 481 AAGGAGCTTGACAGGCGGCCCAAGAGCCTCTCACAACCTTTCTGTGCCCAAGTTCT 540
    |||
Db 481 AAGGAGCTTGACAGGCGGCCCAAGAGCCTCTCACAACCTTTCTGTGCCCAAGTTCT 540
QY 541 TGGGAGTGAATGAATGATCATCATCATGAATCATTTCTGGGCTCTAA 591
    |||
Db 541 TGGGAGTGAATGAATGATCATCATCATGAATCATTTCTGGGCTCTAA 591
```

```
RESULT 2
LOCUS CN304585 729 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000600512783 GRN_PRENED Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN304585
VERSION CN304585.1 GI:47320999
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE 1 (bases 1 to 729)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
            Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
            Nat. Biotechnol. 22 (6), 707-716 (2004)
JOURNAL 15146197
PUBMED Contact: Brandenberger R
COMMENT Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
FEATURES Insert length: 729 Std Error: 0.00.
            Location/Qualifiers
SOURCE 1..729
            /organism="Homo sapiens"
            /mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/tissue_type="embryonic stem cell, retinoic acid and
mitogen-treated hES cell line H7"
/clone_id="GRN_PRENED"
/note="Oligo dT primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic acid and mitogens."
```

ORIGIN

Query Match 99.7%; Score 589.4; DB 7; Length 729;
Best Local Similarity 99.8%; Pred.No.2.1e-135;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAGGAGGCTTGAAGAGAAACACTGTATTTGGAAGAGAGAGAGAGGTGGAG 60
    |||
Db 98 ATGTTGGAGAGCTTGAAGAGAAACACTGTATTTGGAAGAGAGAGAGAGGTGGAG 157
QY 61 TGGAGTCCAGCAGGCGCTTCAAGACTAACAGCAAGCCCTGCTCCGCATCTCCCTAGACAA 120
    |||
Db 158 TGGAGTCCAGCAGGCGCTTCAAGACTAACAGCAAGCCCTGCTCCGCATCTCCCTAGACAA 217
QY 121 GTCCAGCGCAGCCTGGGCCCCGAGACACCCAGCCTCCGAGCATGTCTCATCAATAAC 180
    |||
Db 121 GTCCAGCGCAGCCTGGGCCCCGAGACACCCAGCCTCCGAGCATGTCTCATCAATAAC 180
QY 218 GTCCAGCGCAGCCTGGGCCCCGAGACACCCAGCCTCCGAGCATGTCTCATCAATAAC 277
    |||
Db 218 GTCCAGCGCAGCCTGGGCCCCGAGACACCCAGCCTCCGAGCATGTCTCATCAATAAC 277
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTGCCCTGGCTCCGCCCTGCTGCCCGAG 240
    |||
Db 181 ACCCTCCAACAGCTGCAGGCTGCACTTGCCCTGGCTCCGCCCTGCTGCCCGAG 240
QY 278 ACCCTCCAACAGCTGCAGGCTGCACTTGCCCTGGCTCCGCCCTGCTGCCCGAG 337
    |||
Db 278 ACCCTCCAACAGCTGCAGGCTGCACTTGCCCTGGCTCCGCCCTGCTGCCCGAG 337
QY 241 CCCCTCTTCTGCGGCGAGGAGATTCTCCCTGTCAAGCACCATTGGCTCTATCCTCAGG 300
    |||
Db 241 CCCCTCTTCTGCGGCGAGGAGATTCTCCCTGTCAAGCACCATTGGCTCTATCCTCAGG 300
QY 338 CCCCTCTTCTGCGGCGAGGAGATTCTCCCTGTCAAGCACCATTGGCTCTATCCTCAGG 397
    |||
Db 338 CCCCTCTTCTGCGGCGAGGAGATTCTCCCTGTCAAGCACCATTGGCTCTATCCTCAGG 397
QY 301 GAGCTGACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
    |||
Db 301 GAGCTGACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
QY 398 GAGCTGACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 457
    |||
Db 398 GAGCTGACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 457
QY 361 CTCGAGAATGAAGTGCCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTTGAAGCTCCCGG 420
    |||
Db 361 CTCGAGAATGAAGTGCCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTTGAAGCTCCCGG 420
QY 458 CTCGAGAATGAAGTGCCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTTGAAGCTCCCGG 517
    |||
Db 458 CTCGAGAATGAAGTGCCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTTGAAGCTCCCGG 517
QY 421 TACTTGGGGGACTCTGCGCTGATGACTTCTTTCTGGAACATGACACATCTGCGTAGAA 480
    |||
Db 421 TACTTGGGGGACTCTGCGCTGATGACTTCTTTCTGGAACATGACACATCTGCGTAGAA 480
QY 518 TACTTGGGGGACTCTGCGCTGATGACTTCTTTCTGGAACATGACACATCTGCGTAGAA 577
    |||
Db 518 TACTTGGGGGACTCTGCGCTGATGACTTCTTTCTGGAACATGACACATCTGCGTAGAA 577
QY 481 AAGGAGCTTGACAGGCGGCCCAAGAGCCTCTCACAACCTTTCTGTGCCCAAGTTCT 540
    |||
Db 481 AAGGAGCTTGACAGGCGGCCCAAGAGCCTCTCACAACCTTTCTGTGCCCAAGTTCT 540
QY 578 AAGGAGCTTGACAGGCGGCCCAAGAGCCTCTCACAACCTTTCTGTGCCCAAGTTCT 637
    |||
Db 578 AAGGAGCTTGACAGGCGGCCCAAGAGCCTCTCACAACCTTTCTGTGCCCAAGTTCT 637
QY 541 TGGGAGTGAATGAATGATCATCATCATGAATCATTTCTGGGCTCTAA 591
    |||
Db 638 TGGGAGTGAATGAATGATCATCATCATGAATCATTTCTGGGCTCTAA 688
```

```
RESULT 3
LOCUS CN304584 742 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000597971761 GRN_PRENED Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN304584
VERSION CN304584.1 GI:47320998
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE 1 (bases 1 to 742)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
            Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
```


JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 742 Std Error: 0.00.

FEATURES

Source
1. .742
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, DMSO-treated H9 cell line"
/clone_lib="GRN PREHBP"
/note="Oligo dT primed, full-length enriched cDNA library from DMSO-treated hES cell line H9 (p22) maintained in feeder-free conditions"

ORIGIN

Query Match 99.7%; Score 589.4; DB 7; Length 742;
Best Local Similarity 99.8%; Pred. No. 2.1e-135;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGGGAGGCTTGAAGAGAAACACTGATTGTGAAGAGAGAGAGAGGTGGAG 60
Db 68 ATGTGGGAGGCTTGAAGAGAAACACTGATTGTGAAGAGAGAGAGAGGTGGAG 127
QY 61 TGGAGTCCAGCAGGCGCTTCAGAGCTACAGCAAGCCCTGCTCCGCATCTCCCTAGACAA 120
Db 128 TGGAGTCCAGCAGGCGCTTCAGAGCTACAGCAAGCCCTGCTCCGCATCTCCCTAGACAA 187
QY 121 GTCCAGCGCAGCCTGGGCCCCGAGCAACCAGCCTCCGAGGATGTCTCATCCATAAC 180
Db 188 GTCCAGCGCAGCCTGGGCCCCGAGCAACCAGCCTCCGAGGATGTCTCATCCATAAC 247
QY 181 ACCCTCCAACAGCTGCAGGCTGCATTGCGCTGCTCCGCGCCCTGCCCCCGAG 240
Db 248 ACCCTCCAACAGCTGCAGGCTGCATTGCGCTGCTCCGCGCCCTGCCCCCGAG 307
QY 241 CCCCTCTTCTGGGCGAGGAGATTCTCCCTGTGAGCCACCATTTGGCTTATCTCAGG 300
Db 308 CCCCTCTTCTGGGCGAGGAGATTCTCCCTGTGAGCCACCATTTGGCTTATCTCAGG 367
QY 301 GAGCTGACACCTCCATGATGGAAGTGAAGCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 368 GAGCTGACACCTCCATGATGGAAGTGAAGCCCTCAGAAATCCAGTACTCCCTTGGC 427
QY 361 CTCAGAATGAAGTGCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTTGAGCTCCCG 420
Db 428 CTCAGAATGAAGTGCACCCCAAGCCTGATCCAGTCTTTAGAAAGCTTGAGCTCCCG 487
QY 421 TACTTGGGGACTCTGCGCTGTGATGACTTCTTTGTGACATTGACATCTGCGGTAGAA 480
Db 488 TACTTGGGGACTCTGCGCTGTGATGACTTCTTTGTGACATTGACATCTGCGGTAGAA 547
QY 481 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTTCTGTGCCCCCAGGTTCT 540
Db 548 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTTCTGTGCCCCCAGGTTCT 607
QY 541 TGGGAGTGAATGAATGATCATCATGGAATCATTTCTGGGGTCTCTAA 591
Db 608 TGGGAGTGAATGAATGATCATCATGGAATCATTTCTGGGGTCTCTAA 658

RESULT 4
BMS60308 BMS60308 1038 bp mRNA linear EST 20-FEB-2002
LOCUS
DEFINITION AGENCOURT_6563953 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5744719
5', mRNA sequence.

ACCESSION BMS60308
VERSION BMS60308.1 GI:18804629
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE 1 (bases 1 to 1038)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM12766 row: 1 column: 08
High quality sequence start: 32
High quality sequence stop: 703.

FEATURES

Source
1. .1038
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5744719"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 99.7%; Score 589.4; DB 3; Length 1038;
Best Local Similarity 99.8%; Pred. No. 2.2e-135;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGGGAGGCTTGAAGAGAAACACTGATTGTGAAGAGAGAGAGAGGTGGAG 60
Db 100 ATGTGGGAGGCTTGAAGAGAAACACTGATTGTGAAGAGAGAGAGAGGTGGAG 159
QY 61 TGGAGTCCAGCAGGCGCTTCAGAGCTACAGCAAGCCCTGCTCCGCATCTCCCTAGACAA 120
Db 160 TGGAGTCCAGCAGGCGCTTCAGAGCTACAGCAAGCCCTGCTCCGCATCTCCCTAGACAA 219
QY 121 GTCCAGCGCAGCCTGGGCCCCGAGCAACCAGCCTCCGAGGATGTCTCATCCATAAC 180
Db 220 GTCCAGCGCAGCCTGGGCCCCGAGCAACCAGCCTCCGAGGATGTCTCATCCATAAC 279
QY 181 ACCCTCCAACAGCTGCAGGCTGCATTGCGCTGCTCCGCGCCCTGCCCCCGAG 240
Db 280 ACCCTCCAACAGCTGCAGGCTGCATTGCGCTGCTCCGCGCCCTGCCCCCGAG 339
QY 241 CCCCTCTTCTGGGCGAGGAGATTCTCCCTGTGAGCCACCATTTGGCTTATCTCAGG 300
Db 340 CCCCTCTTCTGGGCGAGGAGATTCTCCCTGTGAGCCACCATTTGGCTTATCTCAGG 399
QY 301 GAGCTGACACCTTCATGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 360
Db 400 GAGCTGACACCTTCATGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 459
QY 361 CTCAGAATGAAGTGCACCCCAAGCCTGATCCAGTCTTTCTTAGAAAGCTGTGAGCTCCCG 420

```
|||||
Db      460 CTCAGAAATGAAGTGCACCCCAAGCCTGATTCAGTCTTTCTTGAAGCTCTGAGCTCCCGG 519
QY      421 TACTGGGGGACTCTGGCTTGATGACTCTTTCTTGACATGACACATCTGGGTAGAA 480
Db      520 TACTGGGGGACTCTGGCTTGATGACTCTTTCTTGACATGACACATCTGGGTAGAA 579
QY      481 AAGAGCTGCACGGGCCCCACAGAGCCTCTCTCAACCTCTTCTGTGCCCCAGTTCT 540
Db      580 AAGAGCTGCACGGGCCCCACAGAGCCTCTCTCAACCTCTTCTGTGCCCCAGTTCT 639
QY      541 TGGAGTGAATGAATGATCATCATCATGAAATCATTTGGGGTCTTAA 591
Db      640 TGGAGTGAATGAATGATCATCATCATGAAATCATTTGGGGTCTTAA 690

RESULT 5
CR605181      1298 bp      mRNA      linear      HTC 21-JUL-2004
LOCUS      full-length cDNA clone CS0DJ012YJ05 of T cells (Jurkat cell line)
DEFINITION      Cot 10-normalized of Homo sapiens (human).
ACCESSION      CR605181
VERSION      CR605181.1 GI:504853988
KEYWORDS      HTC, CNSLT_cDNA.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homnidae; Homo.
REFERENCE      1 (bases 1 to 1298)
AUTHORS      Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished
REMARK      Contact : Feng Liang Email : fliang@lifetech.com URL :
              http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Faraday Avenue
REFERENCE      2 (bases 1 to 1298)
AUTHORS      Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT      1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
              end enriched, double-strand cDNA was digested with Not I and cloned
              into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
              was normalized. Library was constructed by Life Technologies, a
              division of Invitrogen.
FEATURES
  source
    1..1298
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CS0DJ012YJ05"
    /issue_type="T cells (Jurkat cell line) Cot
    10-normalized"
    /plasmid="pCMVSPORT_6"
ORIGIN
Query Match      99.7%; Score 589.4; DB 4; Length 1298;
Best Local Similarity 99.8%; Pred. No. 2.3e-135;
Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ATGAGGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGAGGTGGAG 60
Db      105 ATGCTGGAGAGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGAGGTGGAG 164
QY      61 TGGAGTCCAGCAGGCTTCAAGACTACCAAGAGCCCTGCTCCGCATCTCCCTAGACAA 120
Db      165 TGGAGTCCAGCAGGCTTCAAGACTACCAAGAGCCCTGCTCCGCATCTCCCTAGACAA 224
QY      121 GTCCAGCGCAGCCTGGGCCCCGAGCAACCAGCCTCCGCGAGCATGTCTCATCTAAC 180
Db      225 GTCCAGCGCAGCCTGGGCCCCGAGCAACCAGCCTCCGCGAGCATGTCTCATCTAAC 284
```

```
QY      181 ACCCTCAACAGCTGCAGAGGCTGCACCTTGGCTGTCCCGCCCTGCCCCCGAG 240
Db      285 ACCCTCAACAGCTGCAGAGGCTGCACCTTGGCTGTCCCGCCCTGCCCCCGAG 344
QY      241 CCCCTCTTCTGGGCGAGGAGGATTCTCCCTGTGAGCCACCATTGGCTTATCTCAGG 300
Db      345 CCCCTCTTCTGGGCGAGGAGGATTCTCCCTGTGAGCCACCATTGGCTTATCTCAGG 404
QY      301 GAGCTGACACCTTCATGATGATGAGTGAAGCCCCCTCAGAATCCAGTGAATCCCTTGGC 360
Db      405 GAGCTGACACCTTCATGATGATGAGTGAAGCCCCCTCAGAATCCAGTGAATCCCTTGGC 464
QY      361 CTCAGAATGAAGTCCACCCCAAGCCTGATCCAGTCTTCTTAGAAGCTGTGAGCTCCCG 420
Db      465 CTCAGAATGAAGTCCACCCCAAGCCTGATCCAGTCTTCTTAGAAGCTGTGAGCTCCCG 524
QY      421 TACTTGGGGGACTCTGGCCTTGATGACTTCTTCTTGACATTGACACATCTGGCGTAGAA 480
Db      525 TACTTGGGGGACTCTGGCCTTGATGACTTCTTCTTGACATTGACACATCTGGCGTAGAA 584
QY      481 AAGAGCCTGCACGGGCCCCACAGAGCCTCTCTCAACAACCTCTTCTGTGCCCCAGTTCT 540
Db      585 AAGAGCCTGCACGGGCCCCACAGAGCCTCTCTCAACAACCTCTTCTGTGCCCCAGTTCT 644
QY      541 TGGAGTGAATGAATGATCATCATCATGAAATCATTTCTGGGGTCTTAA 591
Db      645 TGGAGTGAATGAATGATCATCATCATGAAATCATTTCTGGGGTCTTAA 695
```

```
RESULT 6
CR601245      1318 bp      mRNA      linear      HTC 21-JUL-2004
LOCUS      full-length cDNA clone CS0DI057YM14 of Placenta Cot 25-normalized
DEFINITION      of Homo sapiens (human).
ACCESSION      CR601245
VERSION      CR601245.1 GI:50482052
KEYWORDS      HTC, CNSLT_cDNA.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homnidae; Homo.
REFERENCE      1 (bases 1 to 1318)
AUTHORS      Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished
REMARK      Contact : Feng Liang Email : fliang@lifetech.com URL :
              http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Faraday Avenue
REFERENCE      2 (bases 1 to 1318)
AUTHORS      Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT      1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
              end enriched, double-strand cDNA was digested with Not I and cloned
              into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
              was normalized. Library was constructed by Life Technologies, a
              division of Invitrogen.
FEATURES
  source
    1..1318
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CS0DI057YM14"
    /issue_type="Placenta Cot 25-normalized"
    /plasmid="pCMVSPORT_6"
ORIGIN
Query Match      99.7%; Score 589.4; DB 4; Length 1318;
Best Local Similarity 99.8%; Pred. No. 2.3e-135;
```

Matches 590; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
QY	1	ATGAGGAGGAGCTTGAAGAGAAACACTCTGATTGTGAAGAGAGAGAGAGGTGGAG	60	
Db	112	ATGCTGGGAGGCTTGAAGAGAAACACTCTGATTGTGAAGAGAGAGAGAGGTGGAG	171	
QY	61	TGAGTCCAGCAGGCTTTCAGAGCTACCAGCAAGCCCTGCTCCGATCTCCCTAGACAA	120	
Db	172	TGGAGTCCAGCAGGCTTTCAGAGCTACCAGCAAGCCCTGCTCCGATCTCCCTAGACAA	231	
QY	121	GTCAGCGCAGCCTGGGCCCCGAGCACCCAGCCTCCGAGCATGTCTCATCTAAC	180	
Db	232	GTCAGCGCAGCCTGGGCCCCGAGCACCCAGCCTCCGAGCATGTCTCATCTAAC	291	
QY	181	ACCCTCCAACAGCTGCAGGCTGCACTTCGCTGGCTCCCGCCCCCTGCCCCGAG	240	
Db	292	ACCCTCCAACAGCTGCAGGCTGCACTTCGCTGGCTCCCGCCCCCTGCCCCGAG	351	
QY	241	CCCCTCTCTGGGCGAGAGATTCTCCCTGTACGCCACATGGCTCTATCTCAGG	300	
Db	352	CCCCTCTCTGGGCGAGAGATTCTCCCTGTACGCCACATGGCTCTATCTCAGG	411	
QY	301	GAGCTGACACCTCCATGATGGGACTGAGCCCCCTCAGATCCAGTCCCTGGC	360	
Db	412	GAGCTGACACCTCCATGATGGGACTGAGCCCCCTCAGATCCAGTCCCTGGC	471	
QY	361	CTCCAGATGAATGCCACCCGCTGATCCAGTCTTCTTAGAAGCTTGAGCTCCGG	420	
Db	472	CTCCAGATGAATGCCACCCGCTGATCCAGTCTTCTTAGAAGCTTGAGCTCCGG	531	
QY	421	TACTTGGGGGACTCTGGCCTGATGACTTCTTCTTGACATTCACATCTGCGTAGAA	480	
Db	532	TACTTGGGGGACTCTGGCCTGATGACTTCTTCTTGACATTCACATCTGCGTAGAA	591	
QY	481	AAGAGCCTGCACGGGCCCCCAGAGCCTCTCTCAACAACCTTCTGTGCCCCAGTTCT	540	
Db	592	AAGAGCCTGCACGGGCCCCCAGAGCCTCTCTCAACAACCTTCTGTGCCCCAGTTCT	651	
QY	541	TGGAGTGAATGAATGATCATCATCATGAAATCATTTCTGGGGTCTTAA	591	
Db	652	TGGAGTGAATGAATGATCATCATCATGAAATCATTTCTGGGGTCTTAA	702	
RESULT 7				
LOCUS	BU157342	1056 bp	mRNA	linear
DEFINITION	AGENCOURT_6925371 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:5952523			
ACCESSION	BU157342			
VERSION	BU157342.1	GI:22670874		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	1 (bases 1 to 1056)			
AUTHORS	NIH-MGC http://mgs.nci.nih.gov/.			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LLCM2140 row: k column: 20 High quality sequence stop: 587. Location/Qualifiers 1..1056			
FEATURES				
source	1..1056			

/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="IMAGE:5952523"				
/tissue_type="ductal carcinoma, cell line"				
/lab_host="DH10B (phage-resistant)"				
/clone_lib="NIH_MGC_110"				
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAG(G). Library constructed by ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."				
ORIGIN				
Query Match 98.6%; Score 582.6; DB 5; Length 1056;				
Best Local Similarity 99.0%; Pred. No. 1.1e-133;				
Matches 585; Conservative 0; Mismatches 6; Indels 0; Gaps 0;				
QY	1	ATGAGGAGGAGCTTGAAGAGAAACACTCTGATTGTGAAGAGAGAGAGAGGTGGAG	60	
Db	108	ATGCTGGGAGGCTTGAAGAGAAACACTCTGATTGTGAAGAGAGAGAGAGGTGGAG	167	
QY	61	TGAGTCCAGCAGGCTTTCAGAGCTACCAGCAAGCCCTGCTCCGATCTCCCTAGACAA	120	
Db	168	TGGAGTCCAGCAGGCTTTCAGAGCTACCAGCAAGCCCTGCTCCGATCTCCCTAGACAA	227	
QY	121	GTCAGCGCAGCCTGGGCCCCGAGCACCCAGCCTCCGAGCATGTCTCATCTAAC	180	
Db	228	GTCAGCGCAGCCTGGGCCCCGAGCACCCAGCCTCCGAGCATGTCTCATCTAAC	287	
QY	181	ACCCTCCAACAGCTGCAGGCTGCACTTCGCTGGCTCCGCCCTGCCCCCGAG	240	
Db	288	ACCCTCCAACAGCTGCAGGCTGCACTTCGCTGGCTCCGCCCTGCCCCCGAG	347	
QY	241	CCCCTCTCTGGGCGAGAGATTCTCCCTGTAGCCACCATGGCTCTATCTCAGG	300	
Db	348	CCCCTCTCTGGGCGAGAGATTCTCCCTGTAGCCACCATGGCTCTATCTCAGG	407	
QY	301	GAGCTGACACCTCCATGATGGGACTGAGCCCCCTCAGATCCAGTCCCTGGC	360	
Db	408	GAGCTGACACCTCCATGATGGGACTGAGCCCCCTCAGATCCAGTCCCTGGC	467	
QY	361	CTCCAGATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTTGAGCTCCGG	420	
Db	468	CTCCAGATGAAGTGCCACCCAGCCTGATCCAGTCTTCTTAGAAGCTTGAGCTCCGG	527	
QY	421	TACTTGGGGGACTCTGGCCTGATGACTTCTTCTTGACATTCACATCTGCGTAGAA	480	
Db	528	TACTTGGGGGACTCTGGCCTGATGACTTCTTCTTGACATTCACATCTGCGTAGAA	587	
QY	481	AAGAGCCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCAGTTCT	540	
Db	588	AAGAGCCTGCACGGGCCCCCAGAGCCTCTCTCACAACCTCTTGTGCCCCAGTTCT	647	
QY	541	TGGAGTGAATGAATGATCATCATCATGAAATCATTTCTGGGGTCTTAA	591	
Db	648	TGGAGTGAATGAATGATCATCATCATGAAATCATTTCTGGGGTCTTAA	698	
RESULT 8				
LOCUS	CN304590	646 bp	mRNA	linear
DEFINITION	17000600179006 GRN_PRENEM Homo sapiens cDNA 5', mRNA sequence.			
ACCESSION	CN304590			
VERSION	CN304590.1	GI:47321004		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	1 (bases 1 to 646)			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
TITLE	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			

REFERENCE	Hominidae; Homo.				
AUTHORS	1 (bases 1 to 646) Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W.				
TITLE	Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation				
JOURNAL	Nat. Biotechnol. 22 (6), 707-716 (2004)				
PUBMED	15146197				
COMMENT	Contact: Brandenberger R Regenerative Medicine Geron Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbrandenberger@geron.com Insert Length: 646 Std Error: 0.00. Location/Qualifiers				
FEATURES	1..646 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /issue_type="embryonic stem cell, retinoic acid and mitogen-treated hES cell line H7" /clone_id="GRN_PRENBU" /note="oligo dT primed, full-length enriched cDNA library from hES cell line H7 (p29) maintained in feeder-free conditions. Embryoid bodies were generated in the presence of all-trans retinoic and mitogens."				
ORIGIN					
Query Match	98.2%	Score 580.4	DB 7	Length 646	
Best Local Similarity	99.8%	Pred. No. 3.5e-133			
Matches 581, Conservative	0	Mismatches 1	Indels 0	Gaps 0	
QY	1	ATGAGGGAGGCTTGAAGAGAAACACTGTATTGGAGAGAGAGAGAGGTGGAG	60		
DB	65	ATGTGGAGAGCTTGAAGAGAAACACTGTATTGGAGAGAGAGAGAGGTGGAG	124		
QY	61	TGAGTCCAGCAGGCTTCAAGCTACCAAGCCCTGCTCCGATCTCCCTAGACAA	120		
DB	125	TGAGTCCAGCAGGCTTCAAGCTACCAAGCCCTGCTCCGATCTCCCTAGACAA	184		
QY	121	GTCCAGCGCAGCTGGGCCCCCGACACCCAGCCTCCGAGGATGTCCTCATCAATAC	180		
DB	185	GTCCAGCGCAGCTGGGCCCCCGACACCCAGCCTCCGAGGATGTCCTCATCAATAC	244		
QY	181	ACCCTCCAACAGCTGCAGGCTGCACTTGCCCTGCTCCGCCCCCTGCCCCCGAG	240		
DB	245	ACCCTCCAACAGCTGCAGGCTGCACTTGCCCTGCTCCGCCCCCTGCCCCCGAG	304		
QY	241	CCCCCTTCTCTGGGCGAGAGGATTCTCCCTGTCAAGCCACCAATTGGCTTATCTCAGG	300		
DB	305	CCCCCTTCTCTGGGCGAGAGGATTCTCCCTGTCAAGCCACCAATTGGCTTATCTCAGG	364		
QY	301	GAGCTGACACCTCCATGATGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGGC	360		
DB	365	GAGCTGACACCTCCATGATGATGGGACTGAGCCCCCTCAGAATCCAGTGACTCCCTTGGC	424		
QY	361	CTCCAGAATGAAGTGGCAACCCCAAGCTGATCCAGTCTTCTTAAAGCTTGAGCTCCCGG	420		
DB	425	CTCCAGAATGAAGTGGCAACCCCAAGCTGATCCAGTCTTCTTAAAGCTTGAGCTCCCGG	484		
QY	421	TACTTGGGGGACTCTGGCCTGATGATCTTTTCTTGGACATTGACACATCTGCGGTAGAA	480		
DB	485	TACTTGGGGGACTCTGGCCTGATGATCTTTTCTTGGACATTGACACATCTGCGGTAGAA	544		
QY	481	AAGAGCCTGCACGGGCCCCCACCAGAGCTCTCTACAACTCTTCTGTGCCCCAGGTTCT	540		
DB	545	AAGAGCCTGCACGGGCCCCCACCAGAGCTCTCTACAACTCTTCTGTGCCCCAGGTTCT	604		
QY	541	TGGAGTGAATGAATGATGATCAATCATGGAATCATTTCTG 582			

```

Db          605 TGGGAGTGGAAATGAATGATCATCATCATGGAATCAATTCTG 646

RESULT 9
BX398356
LOCUS
DEFINITION
BX398356 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CSODI057YM14 5-PRIME, mRNA sequence.
ACCESSION
BX398356
VERSION
BX398356.2 GI:46846521
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
Organism
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 1017)
Lj,W.B., Gruber,C., Tessee,J. and Polayes,D.
Full-length CDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gl:30613656.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and BcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9131.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CSODI057BG07QPI&c=9131.r.
Location/Qualifiers
1..1017
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI057YM14"
/clisue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and BcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match          97.4%; Score 575.8; DB 5; Length 1017;
Best Local Similarity 99.5%; Pred. No. 5.2e-132;
Matches 588; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 1 ATGAGAGGAGGCTTGAAGAGAAACACTCTGATTTGAAGAGAGAGAGAGGTGGAG 60
    |||||
DB 112 ATGCTGGGAGGCTTGAAGAGAAACACTCTGATTTGAAGAGAGAGAGAGGTGGAG 171
    |||||

OY 61 TGSAGTCCAGCAGGCGCTTCAAGCTACCAGCAGGCGCTCTCCGATCTCCCTAGACAA 120
    |||||
DB 172 TGAAGTCCAGCAGGCGCTTCAAGCTACCAGCAGGCGCTCTCCGATCTCCCTAGACAA 231
    |||||

OY 121 GTCCAGCGCAGCCTGGGCCCCGAGCACCAGCCTCCGCGAGCATGTCTCATCATTAAC 180
    |||||
DB 232 GTCCAGCGCA-CCTGGGCCCCGAGCACCAGCCTCCGCGAGCATGTCTCATCATTAAC 290
    |||||

OY 181 ACCCTCCAACAGCTGCAGGCTGCACTTGCCTGCTCCGCCCCCTGCCCCCCCCGAG 240
    |||||
DB 291 ACCCTCCAACAGCTGCAGGCTGCACTTGCCTGCTCCGCCCCCTGCCCCCCCCGAG 350
    |||||

OY 241 CCCCTCTTCTCTGGGCGAGAGGATTTCTCCCTGTCAAGCCACCATTTGGCTATCCTCAGG 300
    |||||
DB 351 CCCCTCTTCTCTGGGCGAGAGGATTTCTCCCTGTCAAGCCACCATTTGGCTATCCTCAGG 410
    |||||

OY 301 GAGCTGACACCTTCATGATGAGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCTTGGC 360
    |||||

```

Db 411 GAGCTGGAACACTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTGAATCCCTTGGC 470
Qy 361 CTCGAGAATGAAGTGGCCACCCCAAGCTTGATCCAGTCTTTTGAAGAGCTTGAGCTCCCG 420
Db 471 CTCGAGAATGAAGTGGCCACCCCAAGCTTGATCCAGTCTTTTGAAGAGCTTGAGCTCCCG 530
Qy 421 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATGACACATCTGCGGTAGAA 480
Db 531 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATGACACATCTGCGGTAGAA 590
Qy 481 AAGGAGCTTGACAGCGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCAGGTTCT 540
Db 591 AAGGAGCCTGCACGCGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCAGGTTCT 650
Qy 541 TGGGAGTGAATGAATGACTGATCATCATGGAATCATTTGGGGTCTTAA 591
Db 651 TGGGAGTGAATGAATGACTGATCATCATGGAATCATTTGGGGTCTTAA 701

RESULT 10

LOCUS CN304594 712 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000532611443 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN304594
VERSION CN304594.1 GI:47321008
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 712)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W.
TITLE Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert length: 712 Std Error: 0.00.

FEATURES

source
1..712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="Oligo dT primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN

Query Match 97.3%; Score 575.2; DB 7; Length 712;
Best Local Similarity 99.3%; Pred. No. 6.9e-132;
Matches 588; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 1 ATGGAGGAGGCTTGAAGAGAAACACTCTGATTGTGAAGAGAGAGAGAGAGGTGGAG 60
Db 115 ATGTTGGAGGCTTGAAGAGAAACACTCTGATTGTGAAGAGAGAGAGAGAGGTGGAG 174
Qy 61 TGGAGTCCAGCAGGCTTCAAGACTACCAAGCAGCCCTGCTCCGCATCTCCCTAGACAAA 120
Db 175 TGGAGTCCAGCAGGCTTCAAGACTACCAAGCAGCCCTGCTCCGCATCTCCCTAGACAAA 234

Qy 121 GTCCAGCGCAGCCTGGGCCCCCGAGACACCCAGCCTCCGCGAGCATGTCTCATCCATAAC 180
Db 235 GTCCAGCGCAGCCTGGGCCCCCGAGACACCCAGCCTCCGCGAGCATGTCTCATCCATAAC 294
Qy 181 ACCCTCCAACAGCTGCAGGCTGCACTTGGCTGGCTCCCGCCCTGCCCTGCCCGCCGAG 240
Db 295 ACCCTCCAACAGCTGCAGGCTGCACTTGGCTGGCTCCCGCCCTGCCCTGCCCGCCGAG 354
Qy 241 CCCCTCTTCTGGGCGAGGAGATTCTCCCTGTACGCCACCATGGCTCTATCCTCAGG 300
Db 355 CCCCTCTTCTGGGCGAGGAGATTCTCCCTGTACGCCACCATGGCTCTATCCTCAGG 414
Qy 301 GAGCTGGAACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTGAATCCCTTGGC 360
Db 415 GAGCTGGAACACCTCCATGATGGGACTGAGCCCCCTCAGAAATCCAGTGAATCCCTTGGC 474
Qy 361 CTCGAGAATGAAGTGGCCACCCCAAGCCTGATCCAGTCTTTTGAAGAGCTTGAGCTCCCG 420
Db 475 CTCGAGAATGAAGTGGCCACCCCAAGCCTGATCCAGTCTTTTGAAGAGCTTGAGCTCCCG 534
Qy 421 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 480
Db 535 TACTTGGGGGACTCTGGCCTGGATGACTTCTTTCTGGACATTGACACATCTGCGGTAGAA 594
Qy 481 AAGGAGCCTTGACAGCGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCAGGTT-C 539
Db 595 AAGGAGCCTTGACAGCGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCAGGTTCC 654
Qy 540 TGGGAGTGAATGAATGACTGATCATCATGGAATCATTTGGGGTCTTAA 591
Db 655 TGGGAGTGAATGAATGACTGATCATCATGGAATCATTTGGGGTCTTAA 706

RESULT 11

LOCUS BU508982 873 bp mRNA linear EST 12-SEP-2002
DEFINITION AGENCOURT 10095430 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5503515
5', mRNA sequence.
ACCESSION BU508982
VERSION BU508982.1 GI:22815215
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 873)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L14061 row: e column: 20
High quality sequence stop: 605.
location/Qualifiers
1..873
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5503515"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

FEATURES

source
1..873
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5503515"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

ORIGIN

Query Match 96.1%; Score 567.8; DB 5; Length 873;
Best Local Similarity 98.5%; Pred. No. 4.9e-130;
Matches 583; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 ATGAGGGAGGCTTGAAGAGAAACACTCTGATTGTGGAAGAGAGAGAGAGGTGGAG 60
|||||
Db 179 ATGTTGGAGGCTTGAAGAGAAACACTCTGATTGTGGAAGAGAGAGAGAGGTGGAG 238
QY 61 TGGAGTCCAGCAGGCTTTCAGAGCTACCAAGAGCCCTGCTCCGCATCTCCTTAGACAAA 120
|||||
Db 239 TGAAGTCCAGCAGGCTTTCAGAGCTACCAAGAGCCCTGCTCCGCATCTCCTAGACAAA 298
QY 121 GTCCAGCGCAGCCTGGGCCCCGAGACACCCAGCCTCCGAGGCATGTCTCATCATTAAC 180
|||||
Db 299 GTCCAGCGCAGCCTGGGCCCCGAGACACCCAGCCTCCGAGGCATGTCTCATCATTAAC 358
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTCCGCTGGCTCCGCCCCCTGCCCCCGAG 240
|||||
Db 359 ACCCTCCAACAGCTGCAGGCTGCACTTCCGCTGGCTCCGCCCCCTGCCCCCGAG 418
QY 241 CCCCTCTTCTGGGCGAGAGAGATTCTCCCTGTCAAGCCACATTGGCTCTATCTCAGG 300
|||||
Db 419 CCCCTCTTCTGGGCGAGAGAGATTCTCCCTGTCAAGCCACATTGGCTCTATCTCAGG 478
QY 301 GAGCTGACACCTCCATGATGAGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
|||||
Db 479 GAGCTGACACCTCCATGATGAGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 538
QY 361 CTCAGATGAAGTCCACCCAGCCTGATCCAGTCTTCTTGAAGCTCTGAGTCCCGG 420
|||||
Db 539 CTCAGATGAAGTCCACCCAGCCTGATCCAGTCTTCTTGAAGCTCTGAGTCCCGG 598
QY 421 TACTTGGGGGACTCTGGCTGATGACTTCTTCTGACATGACACATCTGCGGTAGAA 480
|||||
Db 599 TACTTGGGGGACTCTGAGCTGATGACTTCTTCTGACATGACACATCTGCGGTAGAA 658
QY 481 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCAGTTTC- 539
|||||
Db 659 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCAGTTTC 718
QY 540 TTGGGAGTGAATGAATGATGATCATCATGAAATCATTTCTGGGGTCTTAA 591
|||||
Db 719 TTGGGAGTGAATGAATGATGATCATCATGAAATCATTTCTGGGGTCTTAA 770

RESULT 12

LOCUS COS79535 970 bp mRNA linear EST 20-JUL-2004
DEFINITION ILLUMIGEN MCO_49981 Katze_MMIL Macaca mulatta cDNA clone
IBIUM:19438 5' similar to Bases 114 to 970 highly similar to human
SERTAD3 (Hs.169138), mRNA sequence.

ACCESSION COS79535
VERSION COS79535
KEYWORDS GI:50410495
SOURCE EST.
ORGANISM Macaca mulatta (rhesus monkey)
Macaca mulatta
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.

REFERENCE 1 (bases 1 to 970)
Magneess,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
Iadonato,S.P.

TITLE Analysis of the Macaca mulatta transcriptome and the sequence
divergence between Macaca and human
JOURNAL Genome Biol. 6 (7), R60 (2005)
PUBMED 15988449

COMMENT Contact: C. Magneess
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400

Fax: 2063780408
Email: cmagneess@illumigen.com
Sequenced on 2004.07.15. 648 Q20 bases. Library Preparation: Prof.
Michael Katze Lab at University of Washington DNA Sequencing:
Illumigen Biosciences Inc. For further information, see
http://www.macaque.org

PCR Primers
FORWARD: CCTCACTAAAGGGAACAAA
BACKWARD: CACTATAGGCGGAATTGGTA
Insert Length: 970 Std Error: 0.00
Plate: CL000412 row: B column: 10
Seq primer: CCTCACTAAAGGGAACAAA
POLYA=No.

FEATURES

source

Location/Qualifiers

1..970

/organism="Macaca mulatta"

/mol_type="mRNA"

/strain="Indian"

/db_xref="taxon:9544"

/clone="IBIUM:19438"

/sex="male"

/dev_stage="adult"

/lab_host="Electromax DH10B"

/clone_1b="Katze_MMIL"

/note="Organ: ileum; Vector: pDONR 222; Site_1: Barg 1;
Site 2: Barg 1; Created from Cloneminer cDNA library
Construction kit (catalog #18249-029)"

ORIGIN

Query Match 94.9%; Score 560.8; DB 7; Length 970;
Best Local Similarity 97.8%; Pred. No. 2.8e-128;
Matches 579; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1 ATGAGGGAGGCTTGAAGAGAAACACTCTGATTGTGGAAGAGAGAGAGAGGTGGAG 60
|||||
Db 121 ATGTTGGAGGCTTGAAGAGAAACACTCTGATTGTGGAAGAGAGAGAGAGGTGGAG 180
QY 61 TGGAGTCCAGCAGGCTTTCAGAGCTACCAAGAGCCCTGCTCCGCATCTCCTTAGACAAA 120
|||||
Db 181 TGGAGTCCAGCAGGCTTTCAGAGCTACCAAGAGCCCTGCTCCGCATCTCCTTAGACAAA 240
QY 121 GTCCAGCGCAGCCTGGGCCCCGAGACACCCAGCCTCCGAGGCATGTCTCATCATTAAC 180
|||||
Db 241 GTCCAGCGCAGCCTGGGCCCCGAGACACCCAGCCTCCGAGGCATGTCTCATCATTAAC 300
QY 241 ACCCTCCAACAGCTGCAGGCTGCACTTCCGCTGGCTCCGCCCCCTGCCCCCGAG 240
|||||
Db 301 ACCCTCCAACAGCTGCAGGCTGCACTTCCGCTGGCTCCGCCCCCTGCCCCCGAG 360
QY 241 CCCCTCTTCTGGGCGAGAGAGATTCTCCCTGTCAAGCCACATTGGCTCTATCTCAGG 300
|||||
Db 361 CCCCTCTTCTGGGCGAGAGAGATTCTCTCTGTCAAGCCACATTGGCTCTATCTCAGG 420
QY 301 GAGCTGACACCTCCATGATGAGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
|||||
Db 421 GAGCTGACACCTCCATGATGAGGAGCTGAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 480
QY 361 CTCAGATGAAGTCCACCCAGCCTGATCCAGTCTTCTTGAAGCTCTGAGCTCCCGG 420
|||||
Db 481 CTCAGATGAAGTCCACCCAGCCTGATCCAGTCTTCTTGAAGCTCTGAGCTCCCGG 540
QY 421 TACTTGGGGGACTCTGGCTGATGACTTCTTCTGACATGACACATCTGCGGTAGAA 480
|||||
Db 541 TACTTGGGGGACTCTGGCTGATGACTTCTTCTGACATGACACATCTGCGGTAGAA 600
QY 481 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCAGTTTC 540
|||||
Db 601 AAGGAGCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCAGTTTC 660
QY 541 TTGGGAGTGG-AATGAATGATCATCATGAAATCATTTCTGGGGTCTTAA 591
|||||
Db 661 TTGGGAGTGAATGAATGATCATCATGAAATCATTTCTGGGGTCTTAA 712

RESULT 13
 LOCUS BG720189 668 bp mRNA linear EST 08-MAY-2001
 DEFINITION 602692233F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4824464 5',
 mRNA sequence.
 ACCESSION BG720189
 VERSION BG720189.1 GI:13999376
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1 (bases 1 to 668)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.

Plate: LLAM10736 Row: a Column: 09
High quality sequence stop: 666.

```

FEATURES
source
Location/Qualifiers
1. 668
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4824464"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.2 kb and normalized to R07 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this is
a NIH MGC Library."
```

ORIGIN

Query Match	94.7%	Score 559.8	DB 2	Length 668
Best Local Similarity	99.5%	Pred. No. 4.6e-128		
Matches 561	Conservative 0	Mismatches 3	Indels 0	Gaps 0
Qy 1	ATGAGGGAGGCTTGAAGAGAACTCTGATTTGGAAGAGAGAGAGAGGTGGAG	60		
Dp 105	ATGTTGGAGGCTTGAAGAGAACTCTGATTTGGAAGAGAGAGAGAGGTGGAG	164		
Qy 61	TGAGTCCAGCAGGCGCTTCAGAGCTACCAAGCGCCTGCTCCGATCTCCCTAGACAA	120		
Dp 165	TGAGTCCAGCAGGCGCTTCAGAGCTACCAAGCAAGCCCTGCTCCGATCTCCCTAGACAA	224		
Qy 121	GTCACGCGCAGCCTGGGCCCCCGAGCAACCCAGCCCTCCGAGGCATGTCTTCATCCATPAAC	180		
Dp 225	GTCACGCGCAGCCTGGGCCCCCGAGCAACCCAGCCCTCCGAGGCATGTCTTCATCCATPAAC	284		
Qy 181	ACCCTCCAAACAGCTGCAGGCTGCACTTGGCCTGGCTCCCGCCCCCTGCCCCCGAG	240		
Dp 285	ACCCTCCAAACAGCTGCAGGCTGCACTTGGCCTGGCTCCCGCCCCCTGCCCCCGAG	344		
Qy 241	CCCCCTTCTCTGGGCGAGAGATTTCTCCCTGCAGCCACCATTTGGCTTATCTCTCAGG	300		
Dp 345	CCCCCTTCTCTGGGCGAGAGATTTCTCCCTGCAGCCACCATTTGGCTTATCTCTCAGG	404		

QY	301	GAGCTGGACACCTCCATGTGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCCCTTGGC	360
Db	405	GAGCTGGACACCTCCATGTGATGGGACTGAGCCCCCTCAGAAATCCAGTGACTCCCCCTTGGG	464
QY	361	CTCCAGAATGAAGTGCCACCCCAAGCCTGATCCAGTCTTCTTAGAAGCTGTGAGCTCCGG	420
Db	465	CTCCAGAATGAAGTGCCACCCCAAGCCTGATCCAGTCTTCTTAGAAGCTGTGAGCTCCGG	524
QY	421	TACTTGGGGACTCTGSCCTGATGACTTCTTCTTGACATTGACACATCGCGGTAGAA	480
Db	525	TACTTGGGGACTCTGSCCTGATGACTTCTTCTTGACATTGACACATCGCGGTAGAA	584
QY	481	AAGGAGCCTGCACGGGGCCCCCACAAGAGCCTCCTCAACAACCTCTTGTGCCCCAGTTCT	540
Db	585	AAGGAGCCTGCACGGGGCCCCCACAAGAGCCTCCTCAACAACCTCTTGTGCCCCAGTTCT	644
QY	541	TGGGAGTGGAAATGAATGTGATCAC	564
Db	645	TGGGAGTGGAAATGAATGTGATCAC	668

RESULT	14
BX362283	
LOCUS	1001 bp mRNA linear EST 08-APR-2004
DEFINITION	BX362283 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED Homo sapiens cDNA clone CS0DJ012YJ05 5-PRIME, mRNA sequence.
ACCESSION	BX362283
VERSION	BX362283.2 GI:46289607
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 1001)	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.	Full-length cDNA libraries and normalization	Unpublished (2001)
				On May 5, 2003 this sequence version replaced gi:30370588.

Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segreff@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dt) primer. Five primer
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9131.r

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?c=CS05DJ012CE03QPI&c=9131.r>.

FEATURES	Location/Qualifiers
source	1. .1001

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ012YJ05"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_1lb="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	94.4%;	Score 558;	DB 5;	Length 1001;
Best Local Similarity	99.1%;	Pred. No. 1.4e-127;		
Matches 581;	Conservative 1;	Mismatches 2;	Indels 2;	Gaps 2;

|||||
Db 105 ATGTGGAGGCTTGAAGAGAAACACTGATTTGAAAGAGAGAGAGAGGTGGAG 164
QY 61 TGAAGTCCAGACAGGCTTCAGAGCTACAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
|||||
Db 165 TGAAGTCCAGACAGGCTTCAGAGCTACAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 224
QY 121 GTCCAGCGCAGCTGGGCCCCGAGCACCAGCCTCCGAGGATGTCTCATCTCAATAAC 180
|||||
Db 225 GTCCAGCGCA-CCTGGGCCCCGAGCACCAGCCTCCGAGGATGTCTCATCTCAATAAC 283
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTCCGCTGGCTCCGCCCCCTGCCCTGCCCGCAG 240
|||||
Db 284 ACCCTCCAACAGCTGCAGGCTGCACTTCCGCTGGCHCCGCCCTGCCCTGCCCGCAG 343
QY 241 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTGACGCCACCATTTGGCTATCTCAGG 300
|||||
Db 344 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTGACGCCACCATTTGGCTATCTCAGG 403
QY 301 GAGCTGGAACCTCCATGATGGAAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
|||||
Db 404 GAGCTGGAACCTCCATGATGGAAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 463
QY 361 CTCCAGATGAAGTGCCACCCAGCCTGATCCAGTCTTTAGAAAGCTCTGAGCTCCCGG 420
|||||
Db 464 CTCCAGATGAAGTGCCACCCAGCCTGATCCAGTCTTTAGAAAGCTCTGAGCTCCCGG 523
QY 421 TACTTGGGGGACTCTGGGCTGATGACTTTCTTTCTGGACATTTGACATCTGCGGTAGAA 480
|||||
Db 524 TACTTGGGGGACTCTGGGCTGATGACTTTCTTTCTGGACATTTGACATCTGCGGTAGAA 583
QY 481 AAGAGCCTGCACGCGGCCCCACAGAGCCTCTCTCAACAACCTTTCTGTGCCCGAGTTCT 540
|||||
Db 584 AAGAGCCTGCACGCGGCCCCACAGAGCCTCTCTCAACAACCTTTCTGTGCCCGAGTTCT 643
QY 541 TG-GAGGTGGAATGAATGATGATCATCATGAAATCATTTCTGGGG 585
|||||
Db 644 TCGGAGTGGAAATGAATGATGATCATCATGAAATCATTTCTGGGG 689

RESULT 15
BMS46306 1045 bp mRNA linear EST 20-FEB-2002
LOCUS BMS46306 1045 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6498625 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5722856
5', mRNA sequence.
ACCESSION BMS46306
VERSION BMS46306.1 GI:18779156
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1045)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bms-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM12709 row: j column: 09
High quality sequence start: 20
High quality sequence stop: 626.

FEATURES
SOURCE
1. 1045
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:5722856"
/lab_host="DH10B"
/clone_1tb="NIH_MGC_125"
/note="Organ: ovary (pool of 3); Vector: PCMV-SPORT6,
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
clacking code 036."

ORIGIN
Query Match 94.2%; Score 557; DB 3; Length 1045;
Best Local Similarity 98.6%; Pred. No. 2.5e-127;
Matches 582; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
QY 1 ATGAGGGAGGCTTGAAGAGAAACACTCTGATTTGGAAGAAGAGAGAGAGGTGGAG 60
Db 208 ATGTGGAGGCTTGAAGAGAAACACTCTGATTTGGAAGAAGAGAGAGAGGTGGAG 267
QY 61 TGAAGTCCAGACAGGCTTCAGAGCTACAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 120
Db 268 TGAAGTCCAGACAGGCTTCAGAGCTACAGCAAGCCCTGCTCCGCATCTCCCTAGACAAA 327
QY 121 GTCCAGCGCAGCTGGGCCCCGAGCACCAGCCTCCGAGGATGTCTCATCTCAATAAC 180
Db 328 GTCCAGCGCAGCTGGGCCCCGAGCACCAGCCTCCGAGGATGTCTCATCTCAATAAC 387
QY 181 ACCCTCCAACAGCTGCAGGCTGCACTTCCGCTGGCTCCGCCCCCTGCCCTGCCCGCAG 240
Db 388 ACCCTCCAACAGCTGCAGGCTGCACTTCCGCTGGCTCCGCCCCCTGCCCTGCCCGCAG 447
QY 241 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTGACGCCACCATTTGGCTCTATCTCAGG 300
Db 448 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTGACGCCACCATTTGGCTCTATCTCAGG 507
QY 301 GAGCTGGAACCTCCATGATGGAAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
Db 508 GAGCTGGAACCTCCATGATGGAAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 567
QY 361 CTCCAGATGAAGTGCCACCCAGCCTGATCCAGTCTTTAGAAAGCTCTGAGCTCCCGG 420
Db 568 CTCCAGATGAAGTGCCACCCAGCCTGATCCAGTCTTTAGAAAGCTCTGAGCTCCCGG 627
QY 421 TACTTGGGGGACTCTGGGCTGATGACTTTCTTTCTGGACATTTGACATCTGCGGTAGAA 480
Db 628 TACTTGGGGGACTCTGGGCTGATGACTTTCTTTCTGGACATTTGACATCTGCGGTAGAA 687
QY 481 AAGAGCCTGCACGCGGCCCCACAGAGCCTCTCACAACCTCTTGTGCCCGCAGTTCT 540
Db 688 AAGAGCCTGCACGCGGCCCCACAGAGCCTCTCACAACCTCTTGTGCCCGCAGTTCT 747
QY 541 TGGAGTGG-AATGAATGATCATCAT-GGAATCATTTCTGGGGTCC 588
Db 748 TGGAGTGGCAATGATCATCATGATCCCTCATGGAAATCATTTCTGGGGTCC 797

Search completed: February 6, 2006, 12:37:27
Job time : 3761 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 11:20:09 ; Search time 172 Seconds
(without alignments)
6107.785 Million cell updates/sec

Title: US-10-069-386A-1
Perfect score: 591
Sequence: 1 atggaaggaggttggaagag.....aaatcattctgggtcctaa 591

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA:*

1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.2	8.3	7218	2	US-08-232-463-14
2	47.4	8.0	2082	3	US-10-076-069-1
3	46.4	7.9	44952	3	US-09-949-016-12197
4	46.4	7.9	44960	3	US-09-949-016-17583
5	46.4	7.9	51552	3	US-09-733-294A-30
6	42.6	7.2	1603	3	US-10-000-489-71
7	41	6.9	1926	3	US-09-249-585A-2
8	41	6.9	1926	3	US-09-410-399-3
9	41	6.9	2580	3	US-09-050-863-2
10	41	6.9	2580	3	US-09-359-081-2
11	41	6.9	5452	2	US-09-130-114-1
12	41	6.9	8705	3	US-09-647-344A-14
13	41	6.9	9600	3	US-08-910-647-1
14	41	6.9	9600	3	US-09-620-925-1
15	41	6.9	10596	2	US-07-884-811-15
16	41	6.9	10596	2	US-07-885-971-15
17	41	6.9	10596	2	US-08-087-783A-15
18	41	6.9	10596	2	US-08-194-088B-15
19	41	6.9	10596	2	US-08-194-087-15
20	41	6.9	10596	6	PCT-US93-04648-15
21	41	6.9	16080	3	US-09-724-566A-48
22	41	6.9	16080	3	US-09-471-669A-48
23	40.8	6.9	50937	3	US-09-428-517-1
24	40.6	6.9	396	2	US-07-872-678A-9

C	25	40.4	6.8	2097	3	US-08-941-445A-10	Sequence 10, Appl
	26	40.2	6.8	2254	3	US-09-016-434-1126	Sequence 1126, Ap
	27	40.2	6.8	2573	2	US-08-884-681-2	Sequence 2, Appli
	28	40.2	6.8	2573	3	US-09-258-643-2	Sequence 2, Appli
	29	40.2	6.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	30	40.2	6.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
C	31	40	6.8	2127	3	US-09-902-540-8358	Sequence 8358, Ap
C	32	40	6.8	7628	3	US-09-902-540-875	Sequence 875, App
C	33	39.6	6.7	1432	3	US-09-620-312D-420	Sequence 420, App
C	34	39.6	6.7	2825	3	US-09-196-390-5	Sequence 5, Appli
C	35	39.6	6.7	2825	3	US-09-952-677-5	Sequence 5, Appli
	36	39.6	6.7	4394	2	US-08-095-734-1	Sequence 1, Appli
	37	39.6	6.7	4394	2	US-08-444-623-1	Sequence 1, Appli
	38	39.6	6.7	4394	3	US-08-471-869-1	Sequence 1, Appli
	39	39.6	6.7	4394	3	US-09-342-563-1	Sequence 1, Appli
	40	39.6	6.7	4394	6	PCT-US94-08267-1	Sequence 1, Appli
	41	39.4	6.7	780	3	US-09-902-540-4017	Sequence 4017, Ap
	42	39.4	6.7	2481	3	US-09-894-998A-35	Sequence 35, Appl
	43	39.4	6.7	2481	3	US-10-237-551-35	Sequence 35, Appl
	44	39.4	6.7	3066	3	US-10-237-551-152	Sequence 152, App
C	45	39.4	6.7	26533	3	US-09-902-540-1199	Sequence 1199, Ap

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367

GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls


```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 17583
; LENGTH: 44960
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(44960)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17583
```

```
Query Match          7.9%; Score 46.4; DB 3; Length 44960;
Best Local Similarity 50.4%; Pred. No. 0.059;
Matches 139; Conservative 0; Mismatches 136; Indels 1; Gaps 1;
```

```
QY      45 GGAGGAGAGGTGGAGTGGAGTCCAGCAGGCCCTTACAGACTACCAAGCCCTGCTCCG 104
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      43405 GGAGGAGAGGGGAGGAGGGGGCCCGCAGCGCTTCCCATCTTCCCGCCAGCAGCAG 43346

QY      105 CATCTCCCTAGACAAAGTCCAGCGCAGCGCTGGGCCCCCGAGCAGCCCAAGCTCCGAGGCA 164
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      43345 CCCCCCCCATCACCCCTGCGCAGCTGGGGCCCCCATCATCTCCAGCAGCCCTGGCCA 43286

QY      165 TGTCTCTATCCATACACCTT-CCAACAGCTGAGGCTGCACTTGCGCTGCTCCGCGCC 223
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      43285 GGCCTTCCATCATCCCCCGCCCAAGCTGGGCCCCCAGCATCTCCCTGCTGCGCTGGGC 43226

QY      224 CTGCCCTGCCCCCGAGGCCCTTCTCTGGGCGAGAGAGATTCTCCCTGTGAGCCACCA 283
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      43225 CTGACTTACTGTATGTCTTCCAGGGTGGGGCTCCCATCTGTATCCCTACCTTCTCT 43166

QY      284 TTGGCTCTATCCTCAGGAGAGCTGACACCTCCATGG 319
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      43165 TCCCTCTCTGCTCCTCAGCATCAGAAACCTCCAGG 43130
```

RESULT 5

```
US-09-733-294A-30/c
; Sequence 30, Application US/09733294A
; Patent No. 6492171
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: William Gaarde
; APPLICANT: Susan M. Freiler
; APPLICANT: Edward V. Wancewicz
; TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION
; FILE REFERENCE: ISPH-0527
; CURRENT APPLICATION NUMBER: US/09/733,294A
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 09/572,423
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 108
; SEQ ID NO 30
; LENGTH: 51552
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)...(11492)
; OTHER INFORMATION: exon 1
; NAME/KEY: intron
; LOCATION: (11493)...(11596)
; OTHER INFORMATION: intron 1
; NAME/KEY: exon
; LOCATION: (11597)...(12950)
; OTHER INFORMATION: exon 2
; NAME/KEY: intron
```

```
; LOCATION: (12951)...(21566)
; OTHER INFORMATION: intron 2
; NAME/KEY: exon
; LOCATION: (21567)...(21762)
; OTHER INFORMATION: exon 3
; NAME/KEY: intron
; LOCATION: (21763)...(23851)
; OTHER INFORMATION: intron 3
; NAME/KEY: exon
; LOCATION: (23852)...(24032)
; OTHER INFORMATION: exon 4
; NAME/KEY: intron
; LOCATION: (24033)...(24719)
; OTHER INFORMATION: intron 4
; NAME/KEY: exon
; LOCATION: (24720)...(24899)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (24900)...(25393)
; OTHER INFORMATION: intron 5
; NAME/KEY: exon
; LOCATION: (25394)...(25549)
; OTHER INFORMATION: exon 6
; NAME/KEY: intron
; LOCATION: (25550)...(30196)
; OTHER INFORMATION: intron 6
; NAME/KEY: exon
; LOCATION: (30195)...(30292)
; OTHER INFORMATION: exon 7
; NAME/KEY: intron
; LOCATION: (30293)...(31272)
; OTHER INFORMATION: intron 7
; NAME/KEY: exon
; LOCATION: (31273)...(31358)
; OTHER INFORMATION: exon 8
; NAME/KEY: intron
; LOCATION: (31359)...(33843)
; OTHER INFORMATION: intron 8
; NAME/KEY: unsure
; LOCATION: 31450
; OTHER INFORMATION: unknown
; NAME/KEY: exon
; LOCATION: (33844)...(33957)
; OTHER INFORMATION: exon 9
; NAME/KEY: intron
; LOCATION: (33958)...(35941)
; OTHER INFORMATION: intron 9
; NAME/KEY: exon
; LOCATION: (35942)...(36013)
; OTHER INFORMATION: exon 10
; NAME/KEY: intron
; LOCATION: (36014)...(37884)
; OTHER INFORMATION: intron 10
; NAME/KEY: exon
; LOCATION: (37885)...(38073)
; OTHER INFORMATION: exon 11
; NAME/KEY: intron
; LOCATION: (38074)...(41874)
; OTHER INFORMATION: intron 11
; NAME/KEY: exon
; LOCATION: (41875)...(42001)
; OTHER INFORMATION: exon 12
; NAME/KEY: intron
; LOCATION: (42002)...(42881)
; OTHER INFORMATION: intron 12
; NAME/KEY: exon
; LOCATION: (42882)...(42943)
; OTHER INFORMATION: exon 13
; NAME/KEY: intron
; LOCATION: (42944)...(46129)
; OTHER INFORMATION: intron 13
; NAME/KEY: exon
; LOCATION: (46130)...(46254)
```

```

; OTHER INFORMATION: exon 14
; NAME/KEY: intron
; LOCATION: (46255)..(47035)
; OTHER INFORMATION: intron 14
; NAME/KEY: exon
; LOCATION: (47036)..(47173)
; OTHER INFORMATION: exon 15
; NAME/KEY: intron
; LOCATION: (47174)..(47709)
; OTHER INFORMATION: intron 15
; NAME/KEY: exon
; LOCATION: (47710)..(50544)
; OTHER INFORMATION: exon 16
;
US-09-733-294A-30

```

Query Match	7.94;	Score 46.4;	DB 3;	Length 51552;
Best Local Similarity	50.4%;	Pred. No. 0.061;		
Matches 139;	Conservative	0;	Mismatches 136;	Indels 1; Gaps 1;
QY	45	GGAGGAGAGTGGAGTGGAGTCCAGCAGGCTTCAGAGCTTACGACGAGCCCTGCTCCG	104	
Db	48820	GGAGGAGGAGGAGGAGGAGCCAGCCAGGCTTCCCATCTTCCCGCCACCCAGACCAG	48761	
QY	105	CATCTCCCTAGACAAGTCCAGCGCAGCTGGGCCCCGAGCACCCAGCTCCGAGGCA	164	
Db	48760	CCCCCCCATCACCCCTGCAACCCAGCTGGGGCCCCCATCATCCTTGCACCCCTGGCCA	48701	
QY	165	TGTCCTCATCCATAACACCTT-CCAAACAGCTGCAGGCTGCATTTGGCTGGCTCCGCCC	223	
Db	48700	GGCCCTCATCATCCCGCGCCGCCCAAGCTGGGGCCCCAGCATCCCTGTGCTCGGGC	48641	
QY	224	CTGCCCTGCCCCCGAGCCCTCTTCTGGGCGAGGAGATTTCCTCCTGTGAGCCACCA	283	
Db	48640	CTGGACTTACTGTATGTCTTCCAGGGTGGGGGCTCCCACTGTCTATCCCTTACCTCCT	48581	
QY	284	TTGGCTTATCCTTCAGGAGCTGCACACCTTCATGG	319	
Db	48580	TCCCCTCCTGCTTCACAGCATCAGAAACCTTCCAGG	48545	

```

RESULT 6
US-10-000-489-71
; Sequence 71, Application US/10000489
; Patent No. 6794363
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US6.DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Jpatent
; SEQ ID NO 71
; LENGTH: 1603
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..7
; NAME/KEY: CDS
;

```

```

; LOCATION: 8..763
; NAME/KEY: 3'UTR
; LOCATION: 764..1603
; NAME/KEY: polyA_signal
; LOCATION: 1562..1567
; NAME/KEY: polyA_site
; LOCATION: 1588..1603
US-10-000-489-71

```

Query Match	7.2%	Score 42.6	DB 3	Length 1603
Best Local Similarity	54.0%	Pred. No. 0.22		
Matches	87	Conservative	0	Mismatches 74, Indels 0, Gaps 0
QY	131	GCCTGGCCCCCGAGCACCACCGCTCCGAGGCATGTCTCATCAATAACACCTCCAC	190	
DB	870	GCACGGAACCCCGCCCACTCCCAATCCCCGCGCCCGCCCTCTCCACCCGTGCTTCCC	929	
QY	191	AGCTGCAGGCTGCACATTGCGCTGGCTCCGCGCCCTGACCTGCCCCCGAGACCCCTCTCC	250	
DB	930	CGCTCCACCCCTCACCTCACCTGCGCCCCCGCCACCCCATCGCGCCCCGGCGGCTGTAT	989	
QY	251	TGGGCGGAGGAGATTCTCCCTGTCAAGCCACCATGGCTCT	291	
DB	990	TGTTTCGCTGGGCTCGGTCGGGCGGCTGTCTCCCTCGGCTCT	1030	

```

RESULT 7
US-09-249-585A-2/c
; Sequence 2, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1926)
; OTHER INFORMATION: coding strand of EBNA-1 DNA
US-09-249-585A-2

```

Query Match	6.9%	Score 41	DB 3	Length 1926
Best Local Similarity	51.4%	Pred. 0.59		
Matches 95	Conservative 0	Mismatches 90	Indels 0	Gaps 0
QY	67	CCAGCAGGCTTTCAGAGCTACGACGAGCCCTGCTCCGCATCTCCCTAGACAAAGTCCAG	126	
Db	1063	CCGGCGGGCTTCACACTACTCTTCGACCCCGGCTCCACTACTCTTCGACCCCGGCTC	100	
QY	127	CGCAGCTTGGGCCCCGAGCACCAGCTCCGAGGCATGTCTCATCATTAACACCTC	186	
Db	1003	CACGTGCTCTCGACCCCGGCTCGACTCTGCTCTGACCCCTCTGCTCTGCCCCCTC	944	
QY	187	CAACAGCTGCAGGCTGCACCTTGGCTCCCGCCCTGACCTGCCCCCGAGCCCCCTC	246	
Db	943	CTCTGCTCTCTGCCCCCTCTGCCCCCTCTGCTCTCTGCCCCCTCTGCTCTCTG	884	
QY	247	TTCTT	251	
Db	883	CCCTT	879	

RESULT 8
US-09-410-399-3/c
; Sequence 3, Application US/09410399
; Patent No. 6482587

GENERAL INFORMATION:
APPLICANT: Robertson, Erle S.
APPLICANT: Cotter, Murray A.
TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
TITLE OF INVENTION: to Genomic Host DNA
FILE REFERENCE: UM-03778
CURRENT APPLICATION NUMBER: US/09/410,399
CURRENT FILING DATE: 1999-10-01
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1926
TYPE: DNA
ORGANISM: Epstein-Barr virus
US-09-410-399-3

Query Match 6.9%; Score 41; DB 3; Length 1926;
Best Local Similarity 51.4%; Pred. No. 0.59;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGAGGCTTCAGAGCTACGAGAGCCCTGCTCCGCACTCTCCTAGACAAAGTCCAG 126
DB 1063 CCGGGGGCTCCACTACTCTCTGACCCCGGCTCCACTACTCTCTGACCCCGGCTC 1004

QY 127 CGCAGCTGGGCCCCCGAGCACCAGCTCCGAGCATGTCTCATTCATACACCCCTC 186
DB 1003 CACTGCTCTCTGACCCCGGCTCCACTCTCTGCTCTGCTCCCTCTGCTCTGCTC 944

QY 187 CAACAGCTGAGGCTGCACTTGCGCTGGCTCCCGCCCTGCGCCCTGAGCCCTC 246
DB 943 CTCTGCTCTGCGCCCTCTGCGCCCTCTGCTCTGCTCTGCGCCCTCTGCTCTG 884

QY 247 TTCCT 251
DB 883 CCCCT 879

RESULT 9
US-09-050-863-2/c
Sequence 2, Application US/09050863
Patent No. 614111
GENERAL INFORMATION:
APPLICANT: Lao, Ying
APPLICANT: Hiang, Betty
APPLICANT: Payan, Don
TITLE OF INVENTION: Mammalian Protein Interaction Cloning
TITLE OF INVENTION: System
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Teat, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/050,863
FILING DATE: 30-MAR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 949-8711
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 2580 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-09-050-863-2

Query Match 6.9%; Score 41; DB 3; Length 2580;
Best Local Similarity 51.4%; Pred. No. 0.65;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGAGGCTTCAGAGCTACGAGAGCCCTGCTCCGCACTCTCCTAGACAAAGTCCAG 126
DB 1446 CCGGGGGCTCCACTACTCTCTGACCCCGGCTCCACTACTCTCTGACCCCGGCTC 1387

QY 127 CGCAGCTGGGCCCCCGAGCACCAGCTCCGAGCATGTCTCATTCATACACCCCTC 186
DB 1386 CACTGCTCTCTGACCCCGGCTCCACTCTGCTCTGCTCCCTCTGCTCTGCTC 1327

QY 187 CAACAGCTGAGGCTGCACTTGCGCTGGCTCCCGCCCTGCGCCCTGAGCCCTC 246
DB 1326 CTCTGCTCTGCGCCCTCTGCGCCCTCTGCTCTGCTCTGCGCCCTCTGCTCTG 1267

QY 247 TTCCT 251
DB 1266 CCCCT 1262

RESULT 10
US-09-359-081-2/c
Sequence 2, Application US/09359081
Patent No. 6316223
GENERAL INFORMATION:
APPLICANT: Lao, Ying
APPLICANT: Hiang, Betty
APPLICANT: Payan, Don
TITLE OF INVENTION: Mammalian Protein Interaction Cloning
TITLE OF INVENTION: System
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Teat, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/359,081
FILING DATE: 22-Jul-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/050,863
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 949-8711
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2580 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-359-081-2

Query Match 6.9%; Score 41; DB 3; Length 2580;
Best Local Similarity 51.4%; Pred. No. 0.65;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCTTCAGAGCTACGAGCAAGCCCTGCTCCGCATCTCCCTAGACAAAGTCCAG 126
DB 1446 CCGCGCGGCTCCACTACTCTCTGACCCCGGCTCCACTACTCTCTGACCCCGGCTC 1387
QY 127 CGCAGCCTGGGCCCCGAGACCCAGCCTCCGAGGATGTCCTCATTCATTAACACCTC 186
DB 1386 CACTGCTCTCTGACCCCGGCTCCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1327
QY 187 CAACAGCTGACGCTGCACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 246
DB 1326 CTCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1267
QY 247 TTCCT 251
DB 1266 CCCCT 1262

RESULT 11

US-09-130-114-1
; Sequence 1, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; TITLE OF INVENTION: From Multiple Transfected Episomes
; FILE REFERENCE: 0867/1D903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: VERBA
US-09-130-114-1

Query Match 6.9%; Score 41; DB 2; Length 5452;
Best Local Similarity 51.4%; Pred. No. 0.8;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCTTCAGAGCTACGAGCAAGCCCTGCTCCGCATCTCCCTAGACAAAGTCCAG 126
DB 1359 CCGCGCGGCTCCACTACTCTCTGACCCCGGCTCCACTACTCTCTGACCCCGGCTC 1418
QY 127 CGCAGCCTGGGCCCCGAGACCCAGCCTCCGAGGATGTCCTCATTCATTAACACCTC 186
DB 1419 CACTGCTCTCTGACCCCGGCTCCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1478
QY 187 CAACAGCTGACGCTGCACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 246
DB 1479 CTCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1538
QY 247 TTCCT 251
DB 1539 CCCCT 1543

RESULT 12

US-09-647-344A-14
; Sequence 14, Application US/09647344A
; Patent No. 6586180
; GENERAL INFORMATION:
; APPLICANT: Ruffner, Duane E.
; APPLICANT: Pierce, Michael L.
; APPLICANT: Chen, Zhidong

; TITLE OF INVENTION: Directed Antisense Libraries
; FILE REFERENCE: T6678.PCT.US
; CURRENT APPLICATION NUMBER: US/09/647,344A
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: PCT/US99/06742
; PRIOR FILING DATE: 1999-03-28
; NUMBER OF SEQ ID NOS: 50
; SEQ ID NO 14
; LENGTH: 8705
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pshuttle
US-09-647-344A-14

Query Match 6.9%; Score 41; DB 3; Length 8705;
Best Local Similarity 51.4%; Pred. No. 0.92;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 67 CCAGCAGGCTTCAGAGCTACGAGCAAGCCCTGCTCCGCATCTCCCTAGACAAAGTCCAG 126
DB 7224 CCGCGCGGCTCCACTACTCTCTGACCCCGGCTCCACTACTCTCTGACCCCGGCTC 7283
QY 127 CGCAGCCTGGGCCCCGAGACCCAGCCTCCGAGGATGTCCTCATTCATTAACACCTC 186
DB 7284 CACTGCTCTCTGACCCCGGCTCCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 7343
QY 187 CAACAGCTGACGCTGCACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 246
DB 7344 CTCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7403
QY 247 TTCCT 251
DB 7404 CCCCT 7408

RESULT 13

US-08-910-647-1/c
; Sequence 1, Application US/08910647
; Patent No. 6251433
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for
; TITLE OF INVENTION: Polynucleotide Delivery
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,647
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fujita, Sharon M.
; REGISTRATION NUMBER: 38,459
; REFERENCE/DOCKET NUMBER: 1218.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

;      TOPOLOGY:  linear
;      MOLECULE TYPE:  DNA (genomic)
US-08-910-647-1

```

Query Match	6.9%	Score 41;	DB 3;	Length 9600;
Best Local Similarity	51.4%;	Pred. No. 0.95;		
Matches 95; Conservative	0;	Mismatches 90;	Indels 0;	Gaps 0;

OY	67	CCAGCAGGCGCTTCAGAGCTACCAAGCAGGCCCTGTCTCCGCATCTCCCTAGACAAGAAGTCCAG	126
Db	1492	CCCCGGGCGCTTCACACTACTCTCTGACC GCCCGGCTTCACACTACTCTGACC CGGCGCTC	1433
OY	127	CGCAGCCTGGGCCCCCGAAGACCCAGCCTCCGACAGGCATGCTCATTCATAACACCCCTC	186
Db	1432	CACCTGCTCTCGAACCCCGGCGCTCACCTCTGCTCTGCCCCCTCTCTGCTC GACCCCTC	1373
OY	187	CAACAGCTGCAGGCTGCACTTCGCGCTGGCTCCCGCCCCCTGCCCCCTGCCCCCGAGCCCCTC	246
Db	1372	CTCCTGCTCTGCCCCCTCTGCTGCCCCCTCTCTGCTCTGCTGCCCCCTCTGCTC CTG	1313
OY	247	TTCCT 251	
Db	1312	CCGCT 1308	

RESULT 14
US-09-620-925-1/c
; Sequence 1, Application US/09620925

GENERAL INFORMATION:

APPLICANT: Zuckermann et al.
TITLE OF INVENTION: Compositions and Methods for Polynucleotide Delivery

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/09/620,925
 FILING DATE: 21-Jul-2000
 CLASSIFICATION: <Unknown>

APPLICATION NUMBER: 08/910,647
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:

NAME: Fujita, Sharon M.
REGISTRATION NUMBER: 38,459
REFERENCE/DOCKET NUMBER: 1218.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-2706
TELEFAX: (510) 655-3542

```

; INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;         LENGTH: 9600 base pairs
;         TYPE: nucleic acid
;         STRANDEDNESS: single
;         TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;     SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-620-925-1

```

Query Match	6.9%	Score 41;	DB 3;	Length 9600;
Best Local Similarity	51.4%	Pred. No. 0.95;		
Matches 95; Conservative	0;	Mismatches 90;	Indels 0;	Gaps 0;

[illegible]

RESULT 15
US-07-884-811-15/c
; Sequence 15, Application US/07884811

GENERAL INFORMATION

APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R
TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd

CITY: South San Francisco
STATE: California

COUNTRY: U
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/884,811
FILING DATE: 19920518

CLASSIFICATION: 530
PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.

REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 755.1

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216

TELEFAX: 415/952-9881
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:

LENGTH: 10596 bases
TYPE: NUCLEIC ACID

STRANDEDNESS: single
TOPOLOGY: linear

US-07-884-811-15

Query Match	6.9%	Score 41;	DB 2;	Length 10596;
Best Local Similarity	51.4%;	Pred. No. 0.98;		
Matches 95;	Conservative 0;	Mismatches 90;	Indels 0;	Gaps 0;

[illegible]

OY 187 CAACAGCTGCAAGGCTGCACTTCGCTGCTCCCCGCTGCCCCGAGCCCTC 246
DB 2906 CTCCTGCTCCTGCCCCCTGCTGCCCCCTCCTGCTCCTGCCCCCTGCTGCCCCCTCCTGCTCCTG 2847
OY 247 TTCCT 251
DB 2846 CCCCCT 2842

Search completed: February 6, 2006, 11:23:17
Job time : 178 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 11:31:10 ; Search time 741 Seconds
(without alignments)
6595.415 Million cell updates/sec

Title: US-10-069-386A-1
Perfect score: 591
Sequence: 1 atggaaggagagcttgaagag.....aatcatctctgggtcctaa 591

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	578.4	97.9	598	US-10-029-386-24898	Sequence 24898, A
2	576.2	97.5	1596	US-10-106-698-1239	Sequence 1239, Ap
3	473	80.0	506	US-09-867-701-6064	Sequence 6064, Ap
4	186.6	31.6	205	US-09-867-701-6239	Sequence 6239, Ap
5	166	28.1	526	US-10-029-386-11178	Sequence 11178, A
6	47.4	8.0	2082	US-10-076-069-1	Sequence 1, Appli
7	46.4	7.9	3179	US-10-840-455-20	Sequence 20, Appl
8	46.4	7.9	25138	US-10-840-455-44	Sequence 44, Appl
9	46.4	7.9	51552	US-09-733-294A-30	Sequence 30, Appl
10	44.8	7.6	440	US-10-184-644-202	Sequence 202, App
11	44.8	7.6	440	US-10-184-634-202	Sequence 202, App
12	44.8	7.6	440	US-10-063-685-52	Sequence 52, Appl
13	44.8	7.6	594	US-10-123-155-10	Sequence 10, Appl
14	44.8	7.6	594	US-10-146-731-10	Sequence 10, Appl
15	44.8	7.6	594	US-10-140-472-10	Sequence 10, Appl
16	44.8	7.6	594	US-10-141-761-10	Sequence 10, Appl
17	44.8	7.6	594	US-10-142-885-10	Sequence 10, Appl
18	44.8	7.6	594	US-10-158-790-10	Sequence 10, Appl
19	44.8	7.6	594	US-10-137-871-10	Sequence 10, Appl
20	44.8	7.6	594	US-10-140-923-10	Sequence 10, Appl
21	44.8	7.6	594	US-10-141-756-10	Sequence 10, Appl
22	44.8	7.6	594	US-10-141-759-10	Sequence 10, Appl
23	44.8	7.6	594	US-10-140-805-10	Sequence 10, Appl

C	24	44.8	7.6	594	6	US-10-140-864-10	Sequence 10, Appl
	25	44.8	7.6	2748	7	US-10-437-963-53188	Sequence 53188, A
	26	42.6	7.2	1603	3	US-09-992-600A-71	Sequence 71, Appl
	27	42.6	7.2	1603	3	US-09-924-340-71	Sequence 71, Appl
	28	42.6	7.2	1603	3	US-09-992-095B-71	Sequence 71, Appl
	29	42.6	7.2	1603	3	US-09-999-570-71	Sequence 71, Appl
	30	42.6	7.2	1603	5	US-10-000-489-71	Sequence 71, Appl
	31	42.6	7.2	1603	5	US-10-000-986-71	Sequence 71, Appl
	32	42.6	7.2	1603	6	US-10-154-678-71	Sequence 71, Appl
	33	42.6	7.2	1603	6	US-10-001-142-71	Sequence 71, Appl
	34	42.6	7.2	1603	8	US-10-838-854-71	Sequence 71, Appl
	35	42.4	7.2	511	8	US-10-029-386-24981	Sequence 24981, A
C	36	42.4	7.2	600	8	US-10-363-345A-39087	Sequence 39087, A
	37	42.4	7.2	600	8	US-10-363-345A-39088	Sequence 39088, A
C	38	42.4	7.2	600	9	US-10-363-483A-39087	Sequence 39087, A
	39	42.4	7.2	600	9	US-10-363-483A-39088	Sequence 39088, A
C	40	42.4	7.2	1650	7	US-10-437-963-27296	Sequence 27296, A
	41	42	7.1	672	7	US-10-437-963-69519	Sequence 69519, A
	42	42	7.1	2580	6	US-10-156-761-3895	Sequence 3895, Ap
	43	42	7.1	9025608	6	US-10-156-761-1	Sequence 1, Appli
	44	41.8	7.1	32329	7	US-10-374-903A-1	Sequence 1, Appli
	45	41.6	7.0	940	7	US-10-437-963-40850	Sequence 40850, A

ALIGNMENTS

RESULT 1
US-10-029-386-24898
; Sequence 24898, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24898
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010271.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
; OTHER INFORMATION: EST HUMAN HIT: BG720189.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: 070622, EVALUE 2.30e-01
; OTHER INFORMATION: NT HIT: g16178011, EVALUE 0.00e+00
US-10-029-386-24898

Query Match	97.9%	Score 578.4;	DB 6;	Length 598;
Best Local Similarity	99.7%	Pred. No. 4.1e-161;		
Matches 590;	Conservative 0;	Mismatches 1;	Indels 1;	Gaps 1;
QY	1	ATGAGGAGAGGCTTGAAGAGAAACACTGTGATTGGAAGAGAGAGAGAGTGGAG	60	
Db	7	ATGATGGAGAGGCTTGAAGAGAAACACTGTGATTGGAAGAGAGAGAGAGTGGAG	66	
QY	61	TGAGATCCAGCAGGCTTCAGAGCTACCAAGCAAGCCCTGCTCCGATCTCCCTAGACAA	120	
Db	67	TGAGATCCAGCAGGCTTCAGAGCTACCAAGCAAGCCCTGCTCCGATCTCCCTAGACAA	126	
QY	121	GTCAGGCGAGCCTGGGCCCCGAGCAACCAAGCCTCCGAGGATGTCTCATCCATAC	180	
Db	127	GTCAGGCGAGCCTGGGCCCCGAGCAACCAAGCCTCCGAGGATGTCTCATCCATAC	186	
QY	181	ACCTTCCAACAGCTGAGGCTGCACTTGCCCTGGCTCCGCCCTGCCCCG	239	
Db	187	ACCTTCCAACAGCTGAGGCTGCACTTGCCCTGGCTCCGCCCTGCCCCG	246	

```
QY 240 GCCCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTCAGCCACCATTTGGCTCTATCTCAG 299
    |||||
Db 247 GCCCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTCAGCCACCATTTGGCTCTATCTCAG 306
QY 300 GGAGCTGACACCTCCATGATGGAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGG 359
    |||||
Db 307 GGAGCTGACACCTCCATGATGGAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGG 366
QY 360 CTTCCAGATGAAGTGCACCCCGAGCTGATCCAGTCTTTAGAGCTGAGCTCCG 419
    |||||
Db 367 CTTCCAGATGAAGTGCACCCCGAGCTGATCCAGTCTTTAGAGCTGAGCTCCG 426
QY 420 GTACTTGGGGGACTCTGGCTGGATGACTTCTTTGTGACATTTGACACATCTGCCGTAGA 479
    |||||
Db 427 GTACTTGGGGGACTCTGGCTGGATGACTTCTTTGTGACATTTGACACATCTGCCGTAGA 486
QY 480 AAAGAGCCTGCACGGGCCCCCAGAGAGCTCTCTCAACCTTTCTGTGCCCCAGGTTT 539
    |||||
Db 487 AAAGAGCCTGCACGGGCCCCCAGAGAGCTCTCTCAACCTTTCTGTGCCCCAGGTTT 546
QY 540 TTGGAGTGAATGAATGATCATCATCATGAATCAATCTTGGGGTCTTAA 591
    |||||
Db 547 TTGGAGTGAATGAATGATCATCATCATGAATCAATCTTGGGGTCTTAA 598
```

RESULT 2

```
US-10-106-698-1239
; Sequence 1239, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 1239
; LENGTH: 1596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc.feature
; LOCATION: (57)..(57)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-1239
```

```
Query Match          97.5%; Score 576.2; DB 5; Length 1596;
Best Local Similarity 99.3%; Pred. No. 2.1e-160;
Matches 587; Conservative 2; Mismatches 0; Indels 2; Gaps 1;
```

```
QY 1 ATGAGGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 60
    |||||
Db 345 ATGAGGGAGGCTTGAAGAGAAACACTGTATTGGAAGAGAGAGAGAGGTGGAG 404
QY 61 TGAGTCCAGACGCGCTTCAAGCTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 120
    |||||
Db 405 TGAGTCCAGACGCGCTTCAAGCTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 464
QY 121 GTCCAGCGCAGCTGGCCCCCGAGACCCAGCCTCCGAGGATGTCTCATCATTAAC 180
    |||||
Db 465 GTCCAGCGC--CCTGGCCCCCGAGACCCAGCCTCCGAGGATGTCTCATCATTAAC 522
QY 181 ACCCTCCAACAGCTGCAGGCTGACTTTCGCTGGCTCCCGCCCTGCCCCCGAG 240
    |||||
Db 523 ACCCTCCAACAGCTGCAGGCTGACTTTCGCTGGCTCCCGCCCTGCCCCCGAG 582
```

```
QY 241 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTCAGCCACCATTTGGCTCTATCTCAG 300
    |||||
Db 583 CCCCTCTTCTGGGCGAGAGGATTTCTCCCTGTCAGCCACCATTTGGCTCTATCTCAG 642
QY 301 GAGCTGACACCTCCATGATGGAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 360
    |||||
Db 643 GAGCTGACACCTCCATGATGGAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGGC 702
QY 361 CTTCCAGATGAAGTGCACCCCGAGCTGATCCAGTCTTTAGAGCTGAGCTCCG 420
    |||||
Db 703 CTTCCAGATGAAGTGCACCCCGAGCTGATCCAGTCTTTAGAGCTGAGCTCCG 762
QY 421 TACTTGGGGGACTCTGGCTGGATGACTTCTTTGTGACATTTGACACATCTGCCGTAGA 480
    |||||
Db 763 TACTTGGGGGACTCTGGCTGGATGACTTCTTTGTGACATTTGACACATCTGCCGTAGA 822
QY 481 AAGAGCCTGCACGGGCCCCCAGAGAGCTCTCTCAACCTTTCTGTGCCCCAGGTTT 540
    |||||
Db 823 AAGAGCCTGCACGGGCCCCCAGAGAGCTCTCTCAACCTTTCTGTGCCCCAGGTTT 882
QY 541 TTGGAGTGAATGAATGATCATCATCATGAATCAATCTTGGGGTCTTAA 591
    |||||
Db 883 TTGGAGTGAATGAATGATCATCATCATGAATCAATCTTGGGGTCTTAA 933
```

RESULT 3

```
US-09-867-701-6064
; Sequence 6064, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6064
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-6064
```

```
Query Match          80.0%; Score 473; DB 3; Length 506;
Best Local Similarity 98.6%; Pred. No. 6.9e-130;
Matches 488; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
```

```
QY 60 GTGAGTCCAGACGCGCTTCAAGCTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 119
    |||||
Db 14 GCGAGTCCAGACGCGCTTCAAGCTACAGCAAGCCCTGCTCCGATCTCCCTAGACAA 73
QY 120 AGTCCAGCGCAGCTGGGCCCCCGAGCACCCAGCCTCCGAGGATGTCTCATCATTA 179
    |||||
Db 74 AGTCCAGCGC--CCTGGGCCCCCGAGCACCCAGCCTCCGAGGATGTCTCATCATTA 131
QY 180 CACCTTCCAACAGCTGCAGGCTGACTTTCGCTGGCTCCCGCCCTGCCCCCGA 239
    |||||
Db 132 CACCTTCCAACAGCTGCAGGATGCACTTTCGCTGGCTCCCGCCCTGCCCCCGA 191
QY 240 GCCCCCTTCTGGGCGAGAGGATTTCTCCCTGTCAGCCACCATTTGGCTCTATCTCAG 299
    |||||
Db 192 GCCCCCTTCTGGGCGAGAGGATTTCTCCCTGTCAGCCACCATTTGGCTCTATCTCAG 251
QY 300 GGAAGTGAACACCTCCATGATGGAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGG 359
    |||||
Db 252 GGAAGTGAACACCTCCATGATGGAGTGAAGCCCCCTCAGAAATCCAGTACTCCCTTGG 311
QY 360 CTTCCAGATGAAGTGCACCCCGAGCCTGATCCAGTCTTTTGAAGCTTGAAGCTTCCG 419
    |||||
```



```
Db      312 CCTCCAGATGAAGTGCCACCCCAAGCTGATCCAGTCTTTAGAAAGCTGTGAGCTCCCG 371
QY      420 GTACTTGGGGGACTCTGGCCCTTGATGACTTCTTTGTGACATTGACACATCTGCGGTGA 479
Db      372 GTACTTGGGGGACTCTGGCCCTTGATGACTTCTTTGTGACATTGACACATCTGCGGTGA 431
QY      480 AAAGAGCCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCCAGGTTTC 539
Db      432 AAAGAGCCTGCACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCCAGGTTTC 491
QY      540 TTGGAGTGAATGA 554
Db      492 TTGGAGTGAATGA 506
```

```
RESULT 4
US-09-867-701-6239
; Sequence 6239, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6239
; LENGTH: 205
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-6239
```

```
Query Match      31.6%; Score 186.6; DB 3; Length 205;
Best Local Similarity 97.1%; Pred. No. 4.8e-45;
Matches 201; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY      63 GAGTCCAGCAGGCTTCAAGCTACAGCAAGCCCTGCTCCGATCTCCCTAGACAAAGT 122
Db      1 GAGTCCAGCAGGCTTCAAGCTACAGCAAGCCCTGCTCCGATCTCCCTAGACAAAGT 60
QY      123 CCAGCGCAGCCTGGGCCCCCGAGCAAGCCTCCGAGGATGTCTCATCCATAACAC 182
Db      61 CCAGCGC--CCTGGGCCCCCGAGCAAGCCTCCGAGGATGTCTCATCCATAACAC 118
QY      183 CCTCCAACAGCTGCAGGCTGACATTTGGCTGCTCCGCCCCCTGCCCCCGAGCC 242
Db      119 CCTCCAACAGCTGCAGGCTGACATTTGGCTGCTCCGCCCCCTGCCCCCGAGCC 178
QY      243 CCTCTTCTGGCGAGAGGATTTCTC 269
Db      179 CCTCTTCTGGCGAGAGGATTTCTC 205
```

```
RESULT 5
US-10-029-386-11178
; Sequence 11178, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11178
```

```
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010271.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
; OTHER INFORMATION: SWISSPROT HIT: Q60963, EVALUE 2.20e+00
; OTHER INFORMATION: NT HIT: g15718683, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: AA426355.1, EVALUE 1.00e-114
US-10-029-386-11178
```

```
Query Match      28.1%; Score 166; DB 6; Length 526;
Best Local Similarity 100.0%; Pred. No. 7e-39;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      426 GGGGACTCTGGCTGATGACTTCTTTGTGACATTGACATCTGCGTAGAAAAGA 485
Db      1 GGGGACTCTGGCTGATGACTTCTTTGTGACATTGACATCTGCGTAGAAAAGA 60
QY      486 GCCTGACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCCAGGTTCTTGGA 545
Db      61 GCCTGACGGGCCCCCAGAGCCTCTCACAACCTCTTGTGCCCCCAGGTTCTTGGA 120
QY      546 GTGAATGACTGATCATCATGAATCATTCCTGGGTCCTAA 591
Db      121 GTGAATGACTGATCATCATGAATCATTCCTGGGTCCTAA 166
```

```
RESULT 6
US-10-076-069-1
; Sequence 1, Application US/10076069
; Publication No. US20020177214A1
; GENERAL INFORMATION:
; APPLICANT: JURECIC, ROLAND
; APPLICANT: NACHTMAN, RONALD
; TITLE OF INVENTION: HEP, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELOPMENT
; FILE REFERENCE: 39532-176599
; CURRENT APPLICATION NUMBER: US/10/076,069
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/268,923
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2082
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (191)..(901)
US-10-076-069-1
```

```
Query Match      8.0%; Score 47.4; DB 5; Length 2082;
Best Local Similarity 55.0%; Pred. No. 0.0012;
Matches 115; Conservative 0; Mismatches 91; Indels 3; Gaps 1;
```

```
QY      8 GAGCCTTGAAGAGAAACACTGTGATTTGAAGAGAGAGAGAGAGTGGAGTGAAGT 67
Db      201 GAGGCTGAAGAGAAATGTGTGACCAAGAGAGAGAGAGGTTTGG--GCAGTG 257
QY      68 CAGCAGCCTTCAAGACTACAGCAAGCCCTGCTCCGCACTCTCCCTAGACAAAGTCCAGC 127
Db      258 TCCCTTCTATAGCTGACGACAGCACTCACTCTGACATGTCCCTTGTCAAGCTCAGC 317
QY      128 GCAGCCTGGGCCCCCGAGACCCAGCCTCCGAGGATGTCTCATATACACCTCC 187
Db      318 TCTGTACATGTCTAGTGAAGCCCAATCTCTGCGGCTCGGTCTCATCGCAACACAGTCC 377
QY      188 AACAGCTGACGGCTGCACTTGCCTGGCT 216
Db      378 GGCAGATCCAGAGAAATGACCAAGAT 406
```

```
RESULT 7
US-10-840-455-20/c
; Sequence 20, Application US/10840455
; Publication No. US20050032094A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; APPLICANT: Hagen, Gustav
; APPLICANT: Wick, Maresa
; APPLICANT: Zubov, Dmitry
; TITLE OF INVENTION: Regulatory DNA Sequences of the Gene for the Human Catalytic
; FILE REFERENCE: Telomerase Subunit, and Their Diagnostic and Therapeutic Use
; FILE REFERENCE: Lea 32 805C1
; CURRENT APPLICATION NUMBER: US/10/840,455
; CURRENT FILING DATE: 2004-05-06
; PRIOR APPLICATION NUMBER: PCT/EP98/08216
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/582,246
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: DE19757984.1
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 3179
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-840-455-20

Query Match          7.9%; Score 46.4; DB 8; Length 3179;
Best Local Similarity 50.4%; Pred. No. 0.0026;
Matches 139; Conservative 0; Mismatches 136; Indels 1; Gaps 1;

QY 45 GGAGAGAGGTGGAGTGAAGTCCAGCAGGCGCTTCAGAGCTACGACGAGCCCTGCTCCG 104
DB 447 GGAGGACAGGGAGGAGGGGCCAGCAGGCTTCCCACTTCCCGCCAGCCAGCAG 388
QY 105 CATCTCCCTAGACAAGTCCAGCGCAGCTGGGCCCCGAGCACCAGCCTCCGAGGCA 164
DB 387 CCCCCCATCACCCCTGCGCACCCAGCTGGGGCCCCCATCATCTGCGCACCTGGCCA 328
QY 165 TGTCTCATCCATAACACCTT-CGAACAGCTGACGGCTGCACTTGCGCTGGCTCCGCCC 223
DB 327 GGGCTTCATCATCCCGCGCGCCCAAGCTGGGGCCCCCAGCATCCCTGCTGCCCTTGGGC 268
QY 224 CTGCCCTGCCCCCGAGCCCTCTTCTTCTGGCGGAGGAGATTCTCCCTGTACGCCACA 283
DB 267 CTGGACTTACTGTATGTCTTCCAGGGTGGGGCTCCCACTGTCTATCCCTTACCTCCT 208
QY 284 TTGGCTCTATCTCAGGAGCTGACACCTCCATGG 319
DB 207 TCCCTCTGCTGCTCAGCAGCATCAGAACTCCAGG 172

RESULT 8
US-10-840-455-44/c
; Sequence 44, Application US/10840455
; Publication No. US20050032094A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; APPLICANT: Hagen, Gustav
; APPLICANT: Wick, Maresa
; APPLICANT: Zubov, Dmitry
; TITLE OF INVENTION: Regulatory DNA Sequences of the Gene for the Human Catalytic
; FILE REFERENCE: Telomerase Subunit, and Their Diagnostic and Therapeutic Use
; FILE REFERENCE: Lea 32 805C1
; CURRENT APPLICATION NUMBER: US/10/840,455
; CURRENT FILING DATE: 2004-05-06
; PRIOR APPLICATION NUMBER: PCT/EP98/08216
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/582,246
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: DE19757984.1
; PRIOR FILING DATE: 1997-12-24
```

```
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
; LENGTH: 25138
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(25138)
; OTHER INFORMATION: n is a, c, g, or t.
US-10-840-455-44

Query Match          7.9%; Score 46.4; DB 8; Length 25138;
Best Local Similarity 50.4%; Pred. No. 0.0035;
Matches 139; Conservative 0; Mismatches 136; Indels 1; Gaps 1;

QY 45 GGAGAGAGGTGGAGTGAAGTCCAGCAGGCGCTTCAGAGCTACGACGAGCCCTGCTCCG 104
DB 22406 GGAGGACAGGGAGGAGGGGCCAGCAGGCTTCCCATCTTCCCGCCAGCCAGCAG 22347
QY 105 CATCTCCCTAGACAAGTCCAGCGCAGCCTGGGCCCCGAGCACCAGCCTCCGAGGCA 164
DB 22346 CCCCCCATCACCCCTGCGCACCCAGCTGGGGCCCCCATCATCTGCGCACCTGGCCA 22287
QY 165 TGTCTCATCCATAACACCTT-CGAACAGCTGACGGCTGCACTTGCGCTGGCTCCGCCC 223
DB 22286 GGGCTTCATCATCCCGCGCGCCCAAGCTGGGGCCCCCAGCATCCCTGCTGCCCTTGGGC 22227
QY 224 CTGCCCTGCCCCCGAGCCCTCTTCTTCTGGCGGAGGAGATTCTCCCTGTACGCCACA 283
DB 22226 CTGGACTTACTGTATGTCTTCCAGGGTGGGGCTCCCACTGTCTATCCCTTACCTCCT 22167
QY 284 TTGGCTCTATCTCAGGAGCTGACACCTCCATGG 319
DB 22166 TCCCTCTGCTGCTCAGCAGCATCAGAACTCCAGG 22131

RESULT 9
US-09-733-294A-30/c
; Sequence 30, Application US/09733294A
; Patent No. US20020045588A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: William Gaarde
; APPLICANT: Susan M. Freier
; APPLICANT: Edward V. Wanciewicz
; TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION
; FILE REFERENCE: ISPH-0527
; CURRENT APPLICATION NUMBER: US/09/733,294A
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 09/572,423
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 108
; SEQ ID NO 30
; LENGTH: 51552
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(11492)
; OTHER INFORMATION: exon 1
; NAME/KEY: intron
; LOCATION: (11493)..(11596)
; OTHER INFORMATION: intron 1
; NAME/KEY: exon
; LOCATION: (11597)..(12950)
; OTHER INFORMATION: exon 2
; NAME/KEY: intron
; LOCATION: (12951)..(21566)
; OTHER INFORMATION: intron 2
; NAME/KEY: exon
; LOCATION: (21567)..(21762)
; OTHER INFORMATION: exon 3
```

```

; NAME/KEY: intron
; LOCATION: (21763)...(23851)
; OTHER INFORMATION: intron 3
; NAME/KEY: exon
; LOCATION: (23852)...(24032)
; OTHER INFORMATION: exon 4
; NAME/KEY: intron
; LOCATION: (24033)...(24719)
; OTHER INFORMATION: intron 4
; NAME/KEY: exon
; LOCATION: (24720)...(24899)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (24900)...(25393)
; OTHER INFORMATION: intron 5
; NAME/KEY: exon
; LOCATION: (25394)...(25549)
; OTHER INFORMATION: exon 6
; NAME/KEY: intron
; LOCATION: (25550)...(30196)
; OTHER INFORMATION: intron 6
; NAME/KEY: exon
; LOCATION: (30195)...(30292)
; OTHER INFORMATION: exon 7
; NAME/KEY: intron
; LOCATION: (30293)...(31272)
; OTHER INFORMATION: intron 7
; NAME/KEY: exon
; LOCATION: (31273)...(31358)
; OTHER INFORMATION: exon 8
; NAME/KEY: intron
; LOCATION: (31359)...(33843)
; OTHER INFORMATION: intron 8
; NAME/KEY: unsure
; LOCATION: 31450
; OTHER INFORMATION: unknown
; NAME/KEY: exon
; LOCATION: (33844)...(33957)
; OTHER INFORMATION: exon 9
; NAME/KEY: intron
; LOCATION: (33958)...(35941)
; OTHER INFORMATION: intron 9
; NAME/KEY: exon
; LOCATION: (35942)...(36013)
; OTHER INFORMATION: exon 10
; NAME/KEY: intron
; LOCATION: (36014)...(37884)
; OTHER INFORMATION: intron 10
; NAME/KEY: exon
; LOCATION: (37885)...(38073)
; OTHER INFORMATION: exon 11
; NAME/KEY: intron
; LOCATION: (38074)...(41874)
; OTHER INFORMATION: intron 11
; NAME/KEY: exon
; LOCATION: (41875)...(42001)
; OTHER INFORMATION: exon 12
; NAME/KEY: intron
; LOCATION: (42002)...(42881)
; OTHER INFORMATION: intron 12
; NAME/KEY: exon
; LOCATION: (42882)...(42943)
; OTHER INFORMATION: exon 13
; NAME/KEY: intron
; LOCATION: (42944)...(46129)
; OTHER INFORMATION: intron 13
; NAME/KEY: exon
; LOCATION: (46130)...(46254)
; OTHER INFORMATION: exon 14
; NAME/KEY: intron
; LOCATION: (46255)...(47035)
; OTHER INFORMATION: intron 14
; NAME/KEY: exon
```

```

; LOCATION: (47036)...(47173)
; OTHER INFORMATION: exon 15
; NAME/KEY: intron
; LOCATION: (47174)...(47709)
; OTHER INFORMATION: intron 15
; NAME/KEY: exon
; LOCATION: (47710)...(50544)
; OTHER INFORMATION: exon 16
; US-09-733-294A-30
```

Query Match 7.9%; Score 46.4; DB 3; Length 51552;
Best Local Similarity 50.4%; Pred. No. 0.0038;
Matches 139; Conservative 0; Mismatches 136; Indels 1; Gaps 1;

```

QY 45 GGAGGAGAGGTGGAGTGAATCCAGCAGGCTTCAGAGCTAACGACGACGACGCTGCTCCG 104
Db 48820 GGAGGAGAGGGGAGGAGGGGGCCAGCCAGGCTTCCCATCTTCCCGCCAGACGACGAG 48761
QY 105 CATCTCCCTAGACAAAGTCCAGGCGCAGCCTGGGCCCCCGAGCAGCCAGCCTCCGAGGCA 164
Db 48760 CCCCCCATCATCACCCCTGCCACCCAGCTGGGGCCCCCATCATCTCCACACCTGGCCA 48701
QY 165 TGTCTCTATCCATTAACACCTT-CCAAAGCTGCAGGCTGCACTTGCCCTGCTCCGCCC 223
Db 48700 GGCCCTCATCATCCCGCCGCCCAAGCTGGGCCCCCAGCATTCCTGCTGCCCTTGGGC 48641
QY 224 CTGCCCTGCCCCCGAGCCCTCTTCTGGGCGAGGAGATTCTCCCTGTACGACCA 283
Db 48640 CTGACTTACTGTATGTCTTCCAGGCTGGGGCTCCCACTGTATCCCTACCTCCT 48581
QY 284 TTGGCTCTATCTTCAGGAGCTGACACCTCCATGG 319
Db 48580 TCCCTCTCTGCTTCACAGCATCAGAAACCTCCAGG 48545
```

RESULT 10
US-10-184-644-202/c
; Sequence 202, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C227
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 202
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-184-644-202

Query Match 7.6%; Score 44.8; DB 5; Length 440;
Best Local Similarity 18.8%; Pred. No. 0.0058;
Matches 58; Conservative 81; Mismatches 169; Indels 0; Gaps 0;

```

QY 127 CGCAGCTGGGCCCCGAGCAGCCAGCCTCCGAGGCATGTCTCATTCATACACCTC 186
Db 389 CSCWSSC.CYRNHSCC..YNC.MS..YKNSBC.CY.C..CS.CSCYT.NC.M.GC.MD 330
QY 187 CAACAGCTGACAGGCTGCACTTGCCCTGGCTCCGCCCCCTGCCCCCGAGGCCCTC 246
```



```
Db      329  CNCCSCCDNCSASCMS.SCSHCYSCSNCSGCCSSCCSSCCSSSCSSSCC 270
QY      247  TTCCTGGCGAGGAGATTTCCTCTGACGCCACCATGGCTCTATCTCAGGAGCTG 306
Db      269  SSSCNNNHNCSSCSCS.SCSGCCSNSSCCSCS..NAGC.N.NSTYBSCRC..TBTC 210
QY      307  GACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTGAATCCCTTGCCCTCCAG 366
Db      209  .ANAC.N..CNCC.CW.TC..NKC.SCTSN.C.RCDBW.AC.CC.NC.C.CC.CC.SC.. 150
QY      367  AATGAAGTGCACCCAGCCTGATCCAGTCTTCTTAGAAGCTTGAGCTCCCGTACTG 426
Db      149  CDCCSA.WTCSDC.BC.WSCYBTHTCDY.BH.T.YC..DCANC.TDIT.CBYNC.THTTC 90
QY      427  GGGGACTC 434
Db      89  .C.B.YBC 82
```

RESULT 11

```
US-10-184-634-202/c
; Sequence 202, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 202
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-634-202
```

Query Match 7.6%; Score 44.8; DB 5; Length 440;
Best Local Similarity 18.8%; Pred. No. 0.0058;
Matches 58; Conservative 81; Mismatches 169; Indels 0; Gaps 0;

```
QY      127  CGCAGCCTGGCCCCGAGCAGCCAGCTCCGAGGAGATGCTCTCATCATACACCTC 186
Db      389  CSCWSSC.CYRNHCSCC..YNC.MS..YKNSBC.CY.C..CS.CSCYT.NC.M.GC.MD 330
QY      187  CAACAGCTGACGCTGACCTTGCTGCTCCGCCCTGCTGCTGCTGCTGCTGCTGCTG 246
Db      329  CNCCSCCDNCSASCMS.SCSHCYSCSNCSGCCSSCCSSCCSSSCSSSCSSSCC 270
QY      247  TTCCTGGCGAGGAGATTTCCTCTGACGCCACCATGGCTCTATCTCAGGAGCTG 306
Db      269  SSSCNNNHNCSSCSCS.SCSGCCSNSSCCSCS..NAGC.N.NSTYBSCRC..TBTC 210
QY      307  GACACCTCCATGATGGAGCTGAGCCCCCTCAGAAATCCAGTGAATCCCTTGCCCTCAG 366
Db      209  .ANAC.N..CNCC.CW.TC..NKC.SCTSN.C.RCDBW.AC.CC.NC.C.CC.CC.SC.. 150
QY      367  AATGAAGTGCACCCAGCCTGATCCAGTCTTCTTAGAAGCTTGAGCTCCCGTACTG 426
Db      149  CDCCSA.WTCSDC.BC.WSCYBTHTCDY.BH.T.YC..DCANC.TDIT.CBYNC.THTTC 90
```

```
QY      427  GGGGACTC 434
Db      89  .C.B.YBC 82
```

RESULT 12

```
US-10-063-685-52/c
; Sequence 52, Application US/10063685
; Publication No. US20030180909A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary B.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,685
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 52
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-685-52
```

Query Match 7.6%; Score 44.8; DB 6; Length 440;
Best Local Similarity 18.8%; Pred. No. 0.0058;
Matches 58; Conservative 81; Mismatches 169; Indels 0; Gaps 0;

```
QY      127  CGCAGCCTGGCCCCGAGCAGCCAGCTCCGAGGAGATGCTCTCATCATACACCTC 186
Db      389  CSCWSSC.CYRNHCSCC..YNC.MS..YKNSBC.CY.C..CS.CSCYT.NC.M.GC.MD 330
QY      187  CAACAGCTGACGCTGACCTTGCTGCTCCGCCCTGCTGCTGCTGCTGCTGCTGCTG 246
Db      329  CNCCSCCDNCSASCMS.SCSHCYSCSNCSGCCSSCCSSCCSSSCSSSCSSSCC 270
QY      247  TTCCTGGCGAGGAGATTTCCTCTGACGCCACCATGGCTCTATCTCAGGAGCTG 306
Db      209  .ANAC.N..CNCC.CW.TC..NKC.SCTSN.C.RCDBW.AC.CC.NC.C.CC.CC.SC.. 150
QY      367  AATGAAGTGCACCCAGCCTGATCCAGTCTTCTTAGAAGCTTGAGCTCCCGTACTG 426
Db      149  CDCCSA.WTCSDC.BC.WSCYBTHTCDY.BH.T.YC..DCANC.TDIT.CBYNC.THTTC 90
QY      427  GGGGACTC 434
Db      89  .C.B.YBC 82
```

RESULT 13

```
US-10-123-155-10/c
; Sequence 10, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary B.
```

```
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-10

Query Match          7.6%; Score 44.8; DB 5; Length 594;
Best Local Similarity 8.7%; Pred.No. 0.0061;
Matches 28; Conservative 135; Mismatches 160; Indels 0; Gaps 0;

QY 79 CAGAGCTACCGAAGCCCTGCTCCGATCTCCCTAGACAAGTCCAGCGAGCCTGGGC 138
   :: : | : : : | : : : : : : : : : : : : : : : : : : : : : :
Db 556 MM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSYA.SYSYS.S.SWS 497

QY 139 CCCCAGACCCAGCCTCCGAGGATGCTCATCCATACACCCCTCCACAGCTGCAG 198
   ::::: : : | : : : | : : : : : : : : : : : : : : : : : : :
Db 496 SYSYSSDDY.CYCCYRHRHCSDSYSYXY.CRCCYTT.SYSRYDCHYSCCSDYYCY 437

QY 199 GCTGCACTTGCGCTGCCGCCGCCCTGCGCTGCCGCCGCCGCCCTTCTTGCGGCGAG 258
   :: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 436 YSYSRYSYSYSWSYSYTDYCSYRCCCYYSYSSSYSSSYSSAYSTSSSSSSSS 377

QY 259 GAGGATTTCTCCCTGTACGCCAATTGCTCTATCTTCAGGAGCTGGACACCTTCATG 318
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 376 YYTSTNYC.T.CC...T..MCAABCSSTTTT...HSCC.SA..A.M.YC.A.S 317

QY 319 GATGGAGCTGAGCCCCCTCAGAACTGAGTACTCCCTGGACCTCCAGAAATGAAGTGCA 378
   :: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 316 YSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.YM.Y..YY.YSYSYCSRKTM.. 257

QY 379 CCCCAGCCTGATCCAGTCTTCTT 401
   : : | : : : : : : : : : : : : : : : : : : : : : :
Db 256 ..TMTDM.T.T..MHMY.KYB.H 234

RESULT 14
US-10-146-731-10/c
; Sequence 10, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
```

```
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C323
; CURRENT APPLICATION NUMBER: US/10/146,731
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-146-731-10

Query Match          7.6%; Score 44.8; DB 6; Length 594;
Best Local Similarity 8.7%; Pred.No. 0.0061;
Matches 28; Conservative 135; Mismatches 160; Indels 0; Gaps 0;

QY 79 CAGAGCTACCGAAGCCCTGCTCCGATCTCCCTAGACAAGTCCAGCGAGCCTGGGC 138
   :: : | : : : | : : : : : : : : : : : : : : : : : : : : : :
Db 556 MM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSYA.SYSYS.S.SWS 497

QY 139 CCCCAGACCCAGCCTCCGAGGATGCTCATCCATACACCCCTCCACAGCTGCAG 198
   ::::: : : | : : : | : : : : : : : : : : : : : : : : : : :
Db 496 SYSYSSDDY.CYCCYRHRHCSDSYSYXY.CRCCYTT.SYSRYDCHYSCCSDYYCY 437

QY 199 GCTGCACTTGCGCTGCCGCCGCCCTGCGCTGCCGCCGCCGCCCTTCTTGCGGCGAG 258
   :: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 436 YSYSRYSYSYSWSYSYTDYCSYRCCCYYSYSSSYSSSYSSAYSTSSSSSSSS 377

QY 259 GAGGATTTCTCCCTGTACGCCAATTGCTCTATCTTCAGGAGCTGGACACCTTCATG 318
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 376 YYTSTNYC.T.CC...T..MCAABCSSTTTT...HSCC.SA..A.M.YC.A.S 317

QY 319 GATGGAGCTGAGCCCCCTCAGAACTGAGTACTCCCTGGACCTCCAGAAATGAAGTGCA 378
   :: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 316 YSYSYS.SSS.S.SYMR.HRA.SHYYTRS..S.MYCY.YM.Y..YY.YSYSYCSRKTM.. 257

QY 379 CCCCAGCCTGATCCAGTCTTCTT 401
   : : | : : : : : : : : : : : : : : : : : : : : : :
Db 256 ..TMTDM.T.T..MHMY.KYB.H 234

RESULT 15
US-10-140-472-10/c
; Sequence 10, Application US/10140472
; Publication No. US2003013888A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C168
; CURRENT APPLICATION NUMBER: US/10/140,472
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
```

```

; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-472-10

```

Query Match	7.6%;	Score 44.8;	DB 6;	Length 594;
Best Local Similarity	8.7%;	Pred. No. 0.0061;		
Matches	28;	Conservative 135;	Mismatches 160;	Indels 0;
			Gaps	0;

QY		79	CAGAGCTACCAGCAAGCCCTGCTCCCGCATCTCCCTAGACAAAGTCCAGGCCGACCTGGGC	138
	:	:	:	:
	:	:	:	:
Dd		556	MM.M.CBT.STT.YA.M.YT.S.S.S.sysSYSYS.S.S.SDSysYA.SysYS.S.SWS	497
	:	:	:	:
	:	:	:	:
QY		139	CCCCGAGCACCAGCCCTCCGAGCATGTCTTCATCCATAACACCCCTCCAAGACTGCAG	198
	:	:	:	:
	:	:	:	:
Dd		496	SYSYSSDDY.CYCCTYYRHCSDSYSYYY.CRCCYT.SYSRYDCHSCCCSDYYCYs	437
	:	:	:	:
	:	:	:	:
QY		199	GCTGCACTTGCCCTGCTCCCGCCCTGCCCCTGCCCCGAGCCCCCTTTCTGGGCGAG	258
	:	:	:	:
	:	:	:	:
Dd		436	YSySRyySYsSYsSWsSYSTDYCsYRcCcYySssSYsSSSAvStSSsssSSs	377
	:	:	:	:
	:	:	:	:
QY		259	GAGGATTCTCCCTGTCAgCACCAATGGCTTATCCTCAGGAgtTGACActTCatG	318
	:	:	:	:
	:	:	:	:
Dd		376	YYTSTNYC.T.CC...T.MCAABCSTTTTTTTTT.HSCC.SA.A.M..YC.A.S	317
	:	:	:	:
	:	:	:	:
QY		319	GATGGGACTGAGCCCCCTCAGAATCCAgtGACTccccCTTGgcCTCCAGAAtGAgtGCCA	378
	:	:	:	:
	:	:	:	:
Dd		316	YSYSYS.SSS.S.sYMR.HRA.shYtRS..S.mycY.yM.y.yy.YSyCSRkTM..	257
	:	:	:	:
	:	:	:	:
QY		379	CCCCAGCCTGATCCAgtCTTCTT	401
	:	:	:	:
	:	:	:	:
Dd		256	..TMTDM.T.T.MMMy.KYB.H	234

Search completed: February 6, 2006, 11:43:45
Job time : 747 secs

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

- Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
- Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 11:34:51 ; Search time 389 Seconds
(without alignments)
1273.268 Million cell updates/sec

Title: US-10-069-386A-1

Perfect score: 591

Sequence: 1 atggaaggaggttggaag.....aatcatctctgggtcctaa 591

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6068529 seqs, 419036697 residues

Total number of hits satisfying chosen parameters: 12137058

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:*
1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
11: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	40.8	6.9	191684	8 US-11-121-086-2	Sequence 2, Appli
c 2	39	6.6	1984	7 US-10-502-893-1	Sequence 1, Appli
c 3	38.2	6.5	415117	7 US-10-995-561-13274	Sequence 13274, A
c 4	37.6	6.4	2910	8 US-11-000-688-628	Sequence 628, App
c 5	37.4	6.3	201	7 US-10-995-561-66759	Sequence 66759, A
c 6	37.4	6.3	18138	7 US-10-995-561-13385	Sequence 13385, A
c 7	37.2	6.3	2139	7 US-10-821-234-596	Sequence 596, App
c 8	37	6.3	2792	8 US-11-136-527-3584	Sequence 3584, App
c 9	37	6.3	3958	8 US-11-136-527-150	Sequence 150, App
c 10	36.8	6.2	153376	8 US-11-121-086-5	Sequence 5, Appli
c 11	36.4	6.2	1161	7 US-10-858-730-170	Sequence 170, App
c 12	36.4	6.2	1529	6 US-10-714-887-113	Sequence 113, App
c 13	36	6.1	185393	8 US-11-121-086-101	Sequence 101, App
c 14	35.8	6.1	45845	8 US-11-153-238-6	Sequence 6, Appli
c 15	35.4	6.0	244	8 US-11-043-752-3912	Sequence 3912, Ap
c 16	35.4	6.0	365	8 US-11-043-752-3906	Sequence 3906, Ap
c 17	35.4	6.0	365	8 US-11-043-752-3909	Sequence 3909, Ap
c 18	35.2	6.0	10968	8 US-11-075-185-35	Sequence 35, Appli
c 19	35.2	6.0	78869	8 US-11-075-185-1	Sequence 1, Appli
c 20	34.8	5.9	1050	7 US-10-750-185-54639	Sequence 54639, A
c 21	34.8	5.9	1050	7 US-10-750-623-54639	Sequence 54639, A
c 22	34.8	5.9	4862	8 US-11-128-061-890	Sequence 890, App

c 23	34.8	5.9	4862	8 US-11-128-049-890	Sequence 890, App
c 24	34.8	5.9	5868	8 US-11-136-527-3354	Sequence 3354, Ap
c 25	34.6	5.9	6432	8 US-11-043-889-3	Sequence 3, Appli
c 26	34.6	5.9	6768	8 US-11-043-889-1	Sequence 1, Appli
c 27	34.2	5.8	4791	8 US-11-000-463-479	Sequence 479, App
c 28	34.2	5.8	150481	8 US-11-112-908-37	Sequence 37, Appli
c 29	34.2	5.8	171162	8 US-11-112-908-38	Sequence 38, Appli
c 30	34	5.8	655	7 US-10-750-185-29911	Sequence 29911, A
c 31	34	5.8	655	7 US-10-750-623-29911	Sequence 29911, A
c 32	34	5.8	2037	8 US-11-136-527-997	Sequence 997, App
c 33	34	5.8	168656	8 US-11-112-908-59	Sequence 59, Appli
c 34	34	5.8	170285	8 US-11-112-908-58	Sequence 58, Appli
c 35	33.8	5.7	1557	7 US-10-996-217A-6	Sequence 6, Appli
c 36	33.8	5.7	3513	7 US-10-858-730-142	Sequence 142, App
c 37	33.8	5.7	5562	8 US-11-180-074-3	Sequence 3, Appli
c 38	33.8	5.7	6990	8 US-11-000-688-609	Sequence 609, App
c 39	33.6	5.7	600	8 US-11-128-061-7073	Sequence 7073, Ap
c 40	33.6	5.7	600	8 US-11-128-049-7073	Sequence 7073, Ap
c 41	33.6	5.7	1182	5 US-09-978-360A-211	Sequence 211, App
c 42	33.6	5.7	1509	7 US-10-750-185-29900	Sequence 29900, A
c 43	33.6	5.7	1509	7 US-10-750-623-29900	Sequence 29900, A
c 44	33.6	5.7	1893	8 US-11-128-061-460	Sequence 460, App
c 45	33.6	5.7	1893	8 US-11-128-049-460	Sequence 460, App

ALIGNMENTS

RESULT 1
US-11-121-086-2/c
; Sequence 2, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121, 086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567, 570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 191684
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-121-086-2

Query Match	6.9%	Score 40.8;	DB 8;	Length 191684;
Best Local Similarity	52.2%;	Pred. No. 1.5;		
Matches 117;	Conservative 0;	Mismatches 102;	Indels 5;	Gaps 1;
QY	75	CCCTCAGAGCTACCAAGACCCCTGCTCCGCATCTCCCTAGACAAAGTCCAGCGACGCT	134	
Db	50337	CCCTCAGAGCTACCAAGACCCCTGCTCCGCATCTCCCTAGACAAAGTCCAGCGACGCT	50278	
QY	135	GGGCCCCCGAGCACCAGCCTCCGAGGATGCTCTCATCATTAACACCCCTCAACAGCT	194	
Db	50277	CCGCGACCTCTCATCTGCTTCAC-----CTTCACACCCCATCAACACCTCATCAGCT	50223	
QY	195	GCAAGCTGACCTTCGCTGCTCCGCCCCCTGCCCCCGAGCCCTCTTCCTGGG	254	
Db	50222	CCACCTTCACCTCCCAACAGCTCTCATGACCTTTCTTTAATCACTCAACCCCTTCCTCTG	50163	
QY	255	CGAGAGGATTTCTCCCTGTCAGCCACCATTTGGCTATCTCA	298	
Db	50162	TTTTCTTCACTCCCTCGATTTCTTCATTATCTCATGACA	50119	
RESULT 2				
US-10-502-893-1				
; Sequence 1, Application US/10502893				


```
; Publication No. US20050255529A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare AG
; APPLICANT: Golz, Stefan
; APPLICANT: Bruggemeier, Ulf
; APPLICANT: Geerts, Andreas
; TITLE OF INVENTION: Diagnostics and Therapeutics for Diseases Associated with a New
; FILE REFERENCE: Human 5HT6 Receptor
; FILE REFERENCE: Lea 35 827
; CURRENT APPLICATION NUMBER: US/10/502, 893
; CURRENT FILING DATE: 2004-07-27
; PRIOR APPLICATION NUMBER: PCT/EP03/000479
; PRIOR FILING DATE: 2003-01-20
; PRIOR APPLICATION NUMBER: EP 02001942.8
; PRIOR FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1984
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-502-893-1
```

```
Query Match          6.6%; Score 39; DB 7; Length 1984;
Best Local Similarity 51.4%; Pred. No. 0.82;
Matches 90; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
```

```
OY 96 CTTGCTCCGCATCTCCCTAGACAAGTCCAGCGCAGCCTGGGCCCCGAGCACCAGCCT 155
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 764 CTTGCTCTGACCGCCTTCGACGTGATGTGTGTCGACGCGCTCCATCTCAACCTGTGCT 823

OY 156 CCGCAGGCATGTCTCTCATCAACAACCTCCCAAGCTGCAAGCTGCACTTGGCCTGGC 215
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 824 CATCAGCCTGACCGCTACTGCTGCTCATCTCTGCGCGCTGCTACAAGCTGCGCATGAC 883

OY 216 TCCCGCCCTGCCCCCTGCCCCCGAGCCCCCTCTTCTTGCGGAGGAGGATTCTCC 270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 884 GCCCCTGCGTGCCCTGCGCCTAGTCTTGCGCGCTGAGCCTCGCCGCTCTCGCC 938
```

```
RESULT 3
US-10-995-561-13274/C
; Sequence 13274, Application US/109955561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995, 561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13274
; LENGTH: 415117
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(415117)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13274
```

```
Query Match          6.5%; Score 38.2; DB 7; Length 415117;
Best Local Similarity 55.7%; Pred. No. 9.6;
Matches 73; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
```

```
OY 125 AGCGCAGCTGGGCCCCGAGCACCAGCCTCCGAGGCATGTCTCATCATACACCC 184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3143 AGCAGAGCCGCGCGCAGCGCCCAACCCGCTCCCACTCCCAAGCGCGAGCGCCGCC 3084

OY 185 TCCAACAGCTGACAGGCTGCACTTGGCTGGCTCCGCGCCCTGCCCCCGAGGCC 244
```

```
Db 3083 GCCCCCGATGCGGCTTCCCCCGCCCTGACACCCCACTGCGCTTCCGCCCCCGCC 3024

OY 245 TCTTCTGCGC 255
    ||||| |||||
Db 3023 GCGCCCTGCGC 3013
```

```
RESULT 4
US-11-000-688-628
; Sequence 628, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, Francois
; APPLICANT: HOUGARTTE, Remi
; APPLICANT: BIRNBAUM, Daniel
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
; FILE REFERENCE: 1423-R-03
; CURRENT APPLICATION NUMBER: US/11/000, 688
; CURRENT FILING DATE: 2004-12-01
; PRIOR APPLICATION NUMBER: US 60/525, 987
; PRIOR FILING DATE: 2003-12-01
; NUMBER OF SEQ ID NOS: 1596
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 628
; LENGTH: 2910
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequences: primer
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2910)
; OTHER INFORMATION: cullin 2(CUL2) gene.
US-11-000-688-628
```

```
Query Match          6.4%; Score 37.6; DB 8; Length 2910;
Best Local Similarity 54.3%; Pred. No. 2.2;
Matches 76; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```

```
OY 130 AGCTGGGCCCCGAGCACCAGCCTCCGAGGCATGTCTCATATACACCTCCAA 189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 54 AGCAGGGGAACGCGAGCTGACAGCGCGCGCGCGCTCCGACCTTCTCGCGC 113

OY 190 CAGCTGACGCTGCACTTGGCTGCTCCGCGCCCTGCCCCCGAGGCCCTTTC 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 114 GGGCTTCTGCTTCTTCACTCTTGGGCTGCTCCCGCTCCCTTGTCTCCCTGCCCCCTTGGC 173

OY 250 CTGGCGAGGAGGATTCTC 269
    ||||| ||||| |||||
Db 174 CTGCTTCTGACAGAAGATTTC 193
```

```
RESULT 5
US-10-995-561-66759
; Sequence 66759, Application US/109955561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995, 561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66759
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-995-561-66759
```

Query Match 6.3%; Score 37.4; DB 7; Length 201;
Best Local Similarity 54.0%; Pred. No. 0.93;
Matches 74; Conservative 1; Mismatches 62; Indels 0; Gaps 0;

```
OY 138 CCCCCGACACCCAGCCTCCGACGCGATGCTCATCCATAACACCCCTCCACAGCTGCA 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 62

OY 198 GCGTGCATTTGCGCTGCGCTCGCTCCCGCCCTGCGCCCTGCGCCCGAGCCCTCTCTGCGCGA 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 63 GCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 122

OY 258 GGAGGATTTCTCCCTGT 274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 123 GCCCGACGCTCTGGCCGT 139
```

RESULT 6

US-10-995-561-13385
; Sequence 13385, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 13385
; LENGTH: 18138
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13385

Query Match 6.3%; Score 37.4; DB 7; Length 18138;
Best Local Similarity 54.0%; Pred. No. 4.9;
Matches 74; Conservative 1; Mismatches 62; Indels 0; Gaps 0;

```
OY 138 CCCCCGACACCCAGCCTCCGACGCGATGCTCATCCATAACACCCCTCCACAGCTGCA 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6274 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6333

OY 198 GCGTGCATTTGCGCTGCGCTCGCTCCCGCCCTGCGCCCTGCGCCCGAGCCCTCTCTGCGCGA 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6334 GCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6393

OY 258 GGAGGATTTCTCCCTGT 274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6394 GCCCGACGCTCTGGCCGT 6410
```

RESULT 7

US-10-821-234-596
; Sequence 596, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 596
; LENGTH: 2139

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-596

Query Match 6.3%; Score 37.2; DB 7; Length 2139;
Best Local Similarity 59.4%; Pred. No. 2.5;
Matches 63; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
OY 218 CCGCCCTGCGCTGCGCCCGCGAGCCCTCTCTCTGCGCGAGGAGGATTTCTCCCTGTCAG 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 408 CCAGCGCGGCTCTCTCTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 467

OY 278 CCACCATGCGCTCTATCTCAGGAGCTGACACCTCATGATGG 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 468 CCCATGCTGCTGGGGGTCCAGGGGCTGAGAGGCTGCTGCTGTTAG 513
```

RESULT 8

US-11-136-527-3584
; Sequence 3584, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays for Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3584
; LENGTH: 2792
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3584

Query Match 6.3%; Score 37; DB 8; Length 2792;
Best Local Similarity 52.2%; Pred. No. 3.2;
Matches 82; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

```
OY 347 TGACTCCCTTGCGCTCCAGATGATGCCACCCAGCCTGATCCAGTCTTCTTAGAAG 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1134 TGCCACCGCTGCTCTCCGGGCTCCCGGACGCCCTGCGCGCTCAGTTCTCCGCTGAG 1193

OY 407 CTCTGAGCTCCCGGTACTTGGGGGACTGTGGCCTGATGATCTTCTTGACATTGACA 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1194 GTCCCTGCGCCAGGTCTTTGATGATTCACCCAGGAGCGCTCTGCTGCTGGAAG 1253

OY 467 CATCTGCGGTAGAAAAGAGCCTGCAGCGGCCCAACC 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1254 CAGCGAACGACAGAGGGGCGCACGCGGAGCCCC 1290
```

RESULT 9

US-11-136-527-150/c
; Sequence 150, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays for Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 150
; LENGTH: 3958
; TYPE: DNA

```
; ORGANISM: Rattus norvegicus
US-11-136-527-150

Query Match          6.3%; Score 37; DB 8; Length 3958;
Best Local Similarity 53.0%; Pred. No. 3.6;
Matches 79; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 79 CAGAGTACCAGCAAGCCCTGCTCCGCATCTCCCTAGACAAGTCCAGCGCAGCCTGGGC 138
    |||||
DB 2072 CAGAACTTCCACTTGCGCCCTTGCCCTCATCTTGGCATCTTTGTCTCATAGCTGTATTGTC 2013

QY 139 CCCCCAGCACCCAGCCTCCGAGGAGCATGTCTCATCCATAACACCCCTCCAACAGCTGCAG 198
    |||||
DB 2012 AAAGCCCCCAGAAAGCCTTGACTGGGAGCGTCTGCTCTTTTCTCTCCTCCAGGTTCTT 1953

QY 199 GCTGCACTTCGCTGGCTCCGCCCCCTGC 227
    |||||
DB 1952 GCAGCACTTGCTGCAGCGGAGCACTTGC 1924

RESULT 10
US-11-121-086-5/c
; Sequence 5, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
; LENGTH: 153376
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-5

Query Match          6.2%; Score 36.8; DB 8; Length 153376;
Best Local Similarity 51.9%; Pred. No. 16;
Matches 83; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 96 CCTGCTCCGCATCTCCCTAGACAAAGTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCT 155
    |||||
DB 16325 CTTCTCTCTCTCTCTCTCTGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 16266

QY 156 CCGCAGGATGTCTCTATCCATAACACCTCCAAACAGCTGAGGCTGCATTCGCTGGC 215
    |||||
DB 16265 CTTGTTCTCTCTCTCTCTCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 16206

QY 216 TCCCGCCCTGCTGCCCCCGAGGCCCTCTTCTGGGC 255
    |||||
DB 16205 CTTCTCTCTCTCTCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGTC 16166

RESULT 11
US-10-858-730-170
; Sequence 170, Application US/10858730
; Publication No. US20050255568A1
; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doten, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.
; APPLICANT: Yorgey, Peter S.
```

```
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858,730
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475,000
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/551,860
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 170
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Streptomyces coelicolor
US-10-858-730-170

Query Match          6.2%; Score 36.4; DB 7; Length 1161;
Best Local Similarity 48.1%; Pred. No. 3.3;
Matches 103; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 96 CTTGCTCCGCATCTCCCTAGACAAAGTCCAGCGCAGCCTGGGCCCCCGAGCACCCAGCCT 155
    |||||
DB 468 CCCGCTGCTCGCATACCGACATCGCGCAGTGGCCCAAGTGCAGCGGAGCCGGGCG 527

QY 156 CCGCAGGATGTCTCTATCCATAACACCTCCAAACAGCTGAGGCTGCATTCGCTGGC 215
    |||||
DB 528 CCGCTCGTGTGACACACCTTCGCCACCCGTACTCCAGCAGCGGCTGGCCCTGCG 587

QY 216 TCCCGCCCTGCTGCCCCCGAGGCCCTCTTCTGAGCGAGGATTTCTCCTGTC 275
    |||||
DB 588 CGCCGACGTGCTGTGACACTCGCTGACCAAGTACATGGGCGGCACTCGACGTGTGG 647

QY 276 AGCCACCATTTGGCTTATCTCTCAGGAGCTGAC 309
    |||||
DB 648 CGCGCGCTGATGTTGGGAGCAGGAGCTGGGC 681

RESULT 12
US-10-714-887-113
; Sequence 113, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPERTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: KUMIMOTO, Neal
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: PINEDA, Omaira
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MB10058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; PRIOR FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 09:53:16 ; Search time 200 Seconds
(without alignments)
430.591 Million cell updates/sec

Title: US-10-069-386A-2
Perfect score: 1033
Sequence: 1 MEGGLKRKHSDLSEEBERWE.....APGSWEWNLDHIMEIILGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*

1:	geneseqp1980s:*
2:	geneseqp1990s:*
3:	geneseqp2000s:*
4:	geneseqp2001s:*
5:	geneseqp2002s:*
6:	geneseqp2003as:*
7:	geneseqp2003bs:*
8:	geneseqp2004s:*
9:	geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	1033	100.0	196	4 AAB35402	Aab35402 Replacat
2	1026	99.3	196	4 AAM93922	Aam93922 Human pol
3	1026	99.3	196	8 ADL32052	Adl32052 Human pro
4	1026	99.3	196	9 ADZ70576	Adz70576 Human pro
5	791	76.6	157	4 AAG74742	Aag74742 Human col
6	645.5	62.5	142	8 ABO60075	Ab060075 Human gen
7	402	38.9	111	8 ABO60074	Ab060074 Human gen
8	216	20.9	236	2 AAY36004	Aay36004 Extended
9	216	20.9	236	8 ADP19312	Adp19312 Human sec
10	213	20.6	236	3 AAY44362	Aay44362 Human cel
11	213	20.6	236	4 AAM93724	Aam93724 Human pol
12	213	20.6	236	4 AAG89292	Aag89292 Human sec
13	213	20.6	236	8 ADL31644	Adl31644 Human pro
14	213	20.6	278	4 AAM25550	Aam25550 Human pro
15	211	20.4	236	2 AAY31829	Aay31829 Human adu
16	205	19.8	222	8 ABO60367	Ab060367 Human gen
17	205	19.8	236	2 AAY02619	Aay02619 Amino aci
18	136	13.2	237	5 AAO22897	Aao22897 Mouse hae
19	125	12.1	314	5 ABP65076	Abp65076 Hypoxia-1
20	125	12.1	314	8 ADR14626	Adr14626 Human NF-
21	125	12.1	314	9 ADX06910	Adx06910 Cyclin-de
22	106	10.3	241	4 AAM39725	Aam39725 Human pol
23	106	10.3	254	3 AAB58258	Aab58258 Lung canc
24	106	10.3	254	4 AAG73682	Aag73682 Human col

25	106	10.3	254	4 AAM41511	Aam41511 Human pol
26	105.5	10.2	578	4 AAM38707	Aam38707 Human pol
27	105.5	10.2	620	7 ADN95361	Adn95361 Human BEC
28	105.5	10.2	620	8 ADP54705	Adp54705 Human PRO
29	105.5	10.2	650	4 AAM38706	Aam38706 Human pol
30	105.5	10.2	685	7 ADE28201	Ade28201 Human MDD
31	105.5	10.2	759	4 AAM40492	Aam40492 Human pol
32	105.5	10.2	759	4 AAM40493	Aam40493 Human pol
33	105.5	10.2	760	7 ADN95128	Adn95128 Human LGC
34	105	10.2	241	5 ABB06375	Abb06375 Human CHD
35	105	10.2	241	5 AAO22898	Aao22898 Human hae
36	105	10.2	241	8 ABM81593	Abm81593 Tumour-as
37	105	10.2	241	9 ADY17410	Ady17410 PRO polyP
38	104.5	10.1	740	4 ABB11713	Abb11713 Human KIA
39	104.5	10.1	879	5 ABP69285	Abp69285 Human pol
40	104	10.1	322	4 ABG22868	Abg22868 Novel hum
41	101.5	9.8	728	7 ADD46841	Add46841 Human pro
42	101.5	9.8	728	7 ADE56294	Ad56294 Human PRO
43	101.5	9.8	728	8 ADJ66562	Adj66562 PI3 kinas
44	98	9.5	562	7 ADB65060	Adb65060 Human pro
45	98	9.5	562	8 ADR58973	Adr58973 Human Btk

ALIGNMENTS

RESULT 1
AAB35402
ID AAB35402 standard, protein; 196 AA.
XX AAB35402;
AC
XX
DT 23-MAY-2001 (first entry)
XX
DE Replication protein A binding transcriptional activator 1 RBT1.
XX
KW RBT1; replication protein A binding transcriptional activator 1; RPA32;
KW gene therapy; apoptosis; cancer; leukaemia.
XX
OS Unidentified.
OS
XX
XX WO200114546-A2.
PN
XX
PD 01-MAR-2001.
PD
XX
PF 17-AUG-2000; 2000WO-CA000948.
PF
XX
PR 19-AUG-1999; 99US-0149472P.
PR
XX
PA (TRAN-) CENT TRANSLATIONAL RES IN CANCER.
XX
XX Alaoui-Jamali MA, Cho JM;
PI
XX WPI; 2001-218447/22.
DR
XX N-PSDB; AAF28052.
DR
XX
PT Novel replication protein A binding transcriptional activator 1 gene,
PT useful for treating neoplastic disorders such as cancer and in gene
PT therapy.
PT
XX
XX
PS Disclosure; Fig 1; 16pp; English.
PS
XX The present invention provides the protein and coding sequences of the
XX replication protein A binding transcriptional activator 1 (RBT1). The
XX protein is capable of inducing apoptosis. The sequences are useful in the
XX gene therapy and other methods of treatment of cancer, including
XX leukaemias. The present sequence is the RBT1 protein
XX
SQ Sequence 196 AA;
SQ
Query Match 100.0%; Score 1033; DB 4; length 196;
Best Local Similarity 100.0%; Pred. No. 1.9e-92;
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MEGGLKRKHSDLSEERERWENSPAGLOSYYQALLRISLDKVQRSIGPRAPSLRRHVLIN 60
DB 1 MEGGLKRKHSDLSEERERWENSPAGLOSYYQALLRISLDKVQRSIGPRAPSLRRHVLIN 60
OY 61 TLQQLQAALRLAPAPALPPEPLFLGEBDFSLSATIGSILRELDTSMGTEPPQNPTPLG 120
DB 61 TLQQLQAALRLAPAPALPPEPLFLGEBDFSLSATIGSILRELDTSMGTEPPQNPTPLG 120
OY 121 LQNEVPPQDPVFLFALSSRYLGDSGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180
DB 121 LQNEVPPQDPVFLFALSSRYLGDSGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180
OY 181 MWMNELDHIMEIILGS 196
DB 181 MWMNELDHIMEIILGS 196

RESULT 2
AAM93922
ID AAM93922 standard; protein; 196 AA.

AC AAM93922;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 4085.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EP130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
DR N-PSDB; AAK94884.

PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.

PS Claim 8; SEQ ID NO 4085; 1380pp + Sequence Listing; English.

XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesized by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a polypeptide encoded by a full length
CC human cDNA of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in CD-ROM
CC format directly from EPO
XX

SQ Sequence 196 AA;

Query Match 99.3%; Score 1026; DB 4; Length 196;
Best Local Similarity 99.5%; Pred. No. 9e-92;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MEGGLKRKHSDLSEERERWENSPAGLOSYYQALLRISLDKVQRSIGPRAPSLRRHVLIN 60
DB 1 MEGGLKRKHSDLSEERERWENSPAGLOSYYQALLRISLDKVQRSIGPRAPSLRRHVLIN 60
OY 61 TLQQLQAALRLAPAPALPPEPLFLGEBDFSLSATIGSILRELDTSMGTEPPQNPTPLG 120
DB 61 TLQQLQAALRLAPAPALPPEPLFLGEBDFSLSATIGSILRELDTSMGTEPPQNPTPLG 120
OY 121 LQNEVPPQDPVFLFALSSRYLGDSGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180
DB 121 LQNEVPPQDPVFLFALSSRYLGDSGLDDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180
OY 181 MWMNELDHIMEIILGS 196
DB 181 MWMNELDHIMEIILGS 196

RESULT 3
ADL32052
ID ADL32052 standard; protein; 196 AA.

AC ADL32052;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human protein encoded by a full length cDNA clone segid 4085.
XX
KW human; medicine; signal transduction; glycoprotein; transcription;
XX oligo-capping method.
XX
OS Homo sapiens.

XX
PN EP1396543-A2.
XX
PD 10-MAR-2004.
XX
PF 07-JUL-2000; 2003EP-00025638.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2004-204755/20.
DR N-PSDB; ADL32051.

PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.

PS Example 1; SEQ ID NO 4085; 1340pp; English.

XX
CC This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polypeptide sequence is a full
CC length human protein of the invention.
XX

SQ Sequence 196 AA;

Query Match 99.3%; Score 1026; DB 8; Length 196;
Best Local Similarity 99.5%; Pred. No. 9e-92;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY		1	MEGGLKRKHSDLSEEEERWWSPPAGLSYQOALLRISLDKVQRSLGPRAPSLRHHVLIHN	60
Dd		1	MVGGLRKSHSDLSEEEERWWSPPAGLSYQOALLRISLDKVQRSLGPRAPSLRHHVLIHN	60
QY		61	TLOQLQAALRLAPAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTETPPONPVTPLG	120
Dd		61	TLOQLQAALRLAPAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTETPPONPVTPLG	120
QY		121	LONEVPQPDPVFVLBALSSRYLGDSDGLDDFFLLDIDTSAVEKEPARAPPEPPHNLFCAPGS	180
Dd		121	LONEVPQPDPVFVLBALSSRYLGDSDGLDDFFLLDIDTSAVEKEPARAPPEPPHNLFCAPGS	180
QY		181	WEWNELDHIMEIILGS	196
Dd		181	WEWNELDHIMEIILGS	196
 RESULT 4 ADZ70576				
ID			ADZ70576 standard; protein; 196 AA.	
XX		AC	ADZ70576;	
XX		DT	30-JUN-2005 (first entry)	
DE			Human protein from lung cancer marker gene RBT1.	
XX				
KM			Tumor marker; lung tumor; cytostatic; neoplasm; expression;	
KM			DNA microarray.	
OS			Homo sapiens.	
XX		PN	WO2005032495-A2.	
Pd			14-APR-2005.	
XX				
PF			01-OCT-2004; 2004WO-US034163.	
PR			03-OCT-2003; 2003US-0508355P.	
XX				
PA			(FARB) BAYER PHARM CORP.	
PI			Taylor I, Pauloski NR, Bigwood D;	
XX				
DR			WPI, 2005-285325/29.	
DR			N-PADB; ADZ70575.	
PT				
PT			Providing a patient diagnosis for lung cancer comprises comparing the	
PT			level of expression of genes or gene products in a biological sample from	
PT			the patient with that from a normal individual.	
XX				
PS			Claim 3; SEQ ID NO 261; 60pp; English.	
CC				
CC			The invention relates to providing a patient diagnosis for lung cancer	
CC			comprising comparing the level of expression of genes or gene products in	
CC			a biological sample from the patient with the level of expression of	
CC			genes or gene products in a biological sample from a normal individual.	
CC			Also included are distinguishing between normal and disease tissues,	
CC			monitoring the response of a patient being treated for lung cancer by	
CC			administering an anti-cancer agent, identifying a compound useful for the	
CC			treatment of lung cancer and an array for distinguishing between normal	
CC			and disease tissues (comprising 2 or more probes corresponding to 2 or	
CC			more genes selected from any of the 200 nucleotide sequences given in the	
CC			specification, or 2 or more polypeptides comprising any of the 200 amino	
CC			acid sequences given in the specification). In providing a patient	
CC			diagnosis for lung cancer, one or more genes are selected from any of the	
CC			200 nucleotide sequences as mentioned in the specification, or one or	
CC			more gene products are polypeptides selected from any of the 20 amino	
CC			acid sequences mentioned in the specification. The methods are useful for	
CC			detecting and treating lung cancer. These may also be used for designing,	
CC			identifying and optimizing therapeutics for cancer. The present sequence	
CC			represents a protein from one of the 200 lung cancer marker genes. Note:	
CC			The sequence data for this patent did not form part of the printed	

```
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 196 AA;

Query Match          99.3%; Score 1026; DB 9; Length 196;
Best Local Similarity 99.5%; Pred. No. 9e-92;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 MEGLKRKHSDEEEERWEMWSAGLSYQQALLRISLDKVQRSLGPRAPSLRRHYLIHN 60
        |||
DB       1 MVGGLRKXKHSDEEEERWEMWSAGLSYQQALLRLSDKVQRSIGPRAPSLRRHYLIHN 60
        |||

QY      61 TLQQLAALRLAPAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTETPPONPVTPLG 120
        |||||
DB       61 TLQQLAALRLAPAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTETPPONPVTPLG 120
        |||||

QY      121 LQNVEVPQPDPVPLEALSSRYLGDSDLDFDLIDITSAVEKEPARAPPBPNNLFCAFGS 180
        |||
DB       121 LQNVEVPQPDPVPLEALSSRYLGDSDLDFDLIDITSAVEKEPARAPPBPNNLFCAFGS 180
        |||

QY      181 WEWNELDHIMEIILGS 196
        |||||
DB       181 WEWNELDHIMEIILGS 196
        |||||

RESULT 5
AAG74742
ID AAG74742 standard; protein; 157 AA.
XX
AC AAG74742;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:5506.
KM Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US026524.
XX
PR 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barabsh SC, Birse CE, Rosen CA;
XX
DR WPI, 2001-235357/24.
DR N-PDSB; AAH34147.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX
PS Claim 11; Page 7098-7099; 9803pp; English.
XX
AAAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
cancer-associated nucleic acid molecules (N) and proteins (P), where the
proteins are collectively known as colon cancer antigens. The colon
cancer antigens have cytosstatic activity and can be used in gene therapy
and vaccine production. N and P may be used in the prevention, diagnosis
cc and treatment of diseases associated with inappropriate P expression. For
example, N and P may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
to affect the patients own production of P. Additionally, N may be used
to supplement the colon cancer-associated PS, by inserting the nucleic acids
to produce the colon cancer-associated PS, by inserting the nucleic acids
```

CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX
SQ Sequence 157 AA;

Query Match 76.6%; Score 791; DB 4; Length 157;
Best Local Similarity 98.0%; Pred. No. 6.2e-69;
Matches 150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 44 SLGPRAPSLRRHVLINHTLQQLQAALRLAPAPALPPRPLFGEDFSLSATIGSILRELD 103
:|||||
DB 5 ALGPRAPSLRRHVLINHTLQQLQAALRLAPAPALPPRPLFGEDFSLSATIGSILRELD 64

QY 104 TSMDGTEPPQNPTPLGLQNEVPPQDPVFLBALSGRYLGDSGLDDFFLDIDTSAVEKEP 163
|||||
DB 65 TSMDGTEPPQNPTPLGLQNEVPPQDPVFLBALSGRYLGDSGLDDFFLDIDTSAVEKEP 124

QY 164 ARAPPEPPHNLFCAPGSWMWNLDHIMEIILGS 196
|||||
DB 125 ARAPPEPXHNLFCAPGSWMWNLDHIMEIILGS 157

RESULT 6
ABO60075
ID ABO60075 standard; protein; 142 AA.
XX
AC ABO60075;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon protein #6309.
XX
KW Human; gene expression; single exon probe; microarray;
XX KW alternative splicing event; genomic alteration.
XX OS Homo sapiens.
XX US2003194704-A1.
XX PN 16-OCT-2003.
XX PD 03-APR-2002; 2002US-00029386.
XX PF 03-APR-2002; 2002US-00029386.
XX PR 03-APR-2002; 2002US-00029386.
XX PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.
DR
XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
XX Claim 45; SEQ ID NO 33709; 80pp; English.
PS
XX
XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human

CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above, a
CC method of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe protein of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
XX
SQ Sequence 142 AA;

Query Match 62.5%; Score 645.5; DB 8; Length 142;
Best Local Similarity 96.8%; Pred. No. 9e-55;
Matches 122; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 71 LAPAPALPPRPLFGEDFSLSATIGSILRELDTSMDGTEPPQNPTPLGLQNEVPPQPD 130
|||||
DB 18 LPPLPC-PPEPLFGEDFSLSATIGSILRELDTSMDGTEPPQNPTPLGLQNEVPPQPD 76

QY 131 PVFLBALSSRYLGDSGLDDFFLDIDTSAVEKEPARAPRPPHNLFCAPGSWMWNLDHIM 190
|||||
DB 77 PVFLBALSSRYLGDSGLDDFFLDIDTSAVEKEPARAPRPPHNLFCAPGSWMWNLDHIM 136

QY 191 EIIILGS 196
|||||
DB 137 EIIILGS 142

RESULT 7
ABO60074
ID ABO60074 standard; protein; 111 AA.
XX
AC ABO60074;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon protein #6308.
XX
KW Human; gene expression; single exon probe; microarray;
XX KW alternative splicing event; genomic alteration.
XX OS Homo sapiens.
XX US2003194704-A1.
XX PN 16-OCT-2003.
XX PD 03-APR-2002; 2002US-00029386.
XX PF 03-APR-2002; 2002US-00029386.
XX PR 03-APR-2002; 2002US-00029386.
XX PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.

XX Penn SG, Rank DR, Hanzel DK;
XX WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.

PS Claim 45; SEQ ID NO 33708; 80pp; English.

XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe protein of the invention. Note: The sequence data for
CC this patent did not form part of the invention. Note: The sequence data for
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX Sequence 111 AA;

Query Match 38.9%; Score 402; DB 8; Length 111;
Best Local Similarity 89.9%; Pred. No. 4.1e-31;
Matches 80; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 MEGGLKRRKHSDLSEERWENSPAGLSYQOALRLISLDKVQRSIGPRAPSLRRHYLIHN 60
Db 3 MWGGLKRRKHSDLSEERWENSPAGLSYQOALRLISLDKVQRSIGPRAPSLRRHYLIHN 62

QY 61 TLQQLQALRLAPAPALPPEPLFLGEEDF 89
Db 63 TLQQLQALRLAPAPALPPEPLPGRGGF 91

RESULT 8
ID AAY36004 standard; protein; 236 AA.

XX AAY36004;
XX
DT 13-SEP-1999 (first entry)
XX
DE Extended human secreted protein sequence, SEQ ID NO. 389.
XX

KW Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW cellular differentiation; immune system regulator; anti-inflammatory;
KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW genetic disease.

OS Homo sapiens.

XX W09931236-A2.

PN 24-JUN-1999.

XX 17-DEC-1998; 98WO-IB002122.

PR 17-DEC-1997; 97US-0069957P.

PR 09-FEB-1998; 98US-0074121P.

PR 13-APR-1998; 98US-0081563P.

PR 10-AUG-1998; 98US-0096116P.

XX (GEST) GENSET.

XX Bougueleret L, Duclert A, Dumas Milne Edwards J;

XX WPI; 1999-385906/32.

XX N-PSDB; AAX97688.

XX New isolated human secreted proteins.

XX Claim 9; Page 330-331; 516pp; English.

XX This sequence is encoded by an extended human secreted protein coding
CC sequence of the invention. The secreted proteins can be used in treating
CC or controlling a variety of human conditions. The secreted proteins may
CC act as cytokines or may affect cellular proliferation or differentiation
CC or may act as immune system regulators, haematopoiesis regulators, tissue
CC growth regulators, regulators of reproductive hormones or cell movement
CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
CC tumour inhibition activity. The DNAs can be used in forensic procedures
CC to identify individuals or in diagnostic procedures to identify
CC individuals having genetic diseases resulting from abnormal expression of
CC the genes corresponding to the extended cDNAs. They are also useful for
CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases

XX Sequence 236 AA;

Query Match 20.9%; Score 216; DB 2; Length 236;
Best Local Similarity 32.6%; Pred. No. 1.6e-12;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRRKHSDLSEEB---ERWENSPAGLSYQO-----LRLISLDKVQRSIGPRAPS 51
Db 5 GLKRRKEEBBKKPLAVDSW-WLDPGHAAVAQAPPAVASSSLFDLSVLKLIHSLQXSXP 63

QY 52 LRRHVLIHNTLQQLQALRLAPAPALPPEPL-----FLGEEDFSLSATIGSILRE 101
Db 64 LRHLVLVXNTLRRIQAS--MAPAALPVPPTPPAAPXVADNLLASSDAALASMAXLLED 121

QY 102 LDTSMDGTEPPQNPTPLGLQNEVPPQPD----PVFLDAL-----SSRYLGDGSLDDEF 152
Db 122 L-SHIGLSQAPQ-----LADEGPPGRSGIXGXPLGALDILGPATGCLDNGLEGLFE 175

QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSNE-----WNELDHIMEIILGS 196
Db 176 DIDTSMYDNEIWPASEGLKPPBD---GPGKEAPDELDEALDYLMADVVG 225

RESULT 9
ID ADP19312 standard; protein; 236 AA.

XX ADP19312;
XX

[illegible]

		RESULT 10
ID	AAY44362	standard; protein; 236 AA.
AC	AAY44362;	
DT	14-MAR-2000	(first entry)
DE	Human cell cycle regulation protein-3.	
KM	CECRP-3; cell cycle regulation protein-3; cell proliferation; cell proliferative disease; cancer; atherosclerosis; cirrhosis; hepatitis; psoriasis; immune system disorder; allergy; asthma; acquired immune deficiency syndrome; Crohn's disease; Blast method; rheumatoid arthritis; gene therapy; chromosomal mapping.	
OS	Homo sapiens.	
FT	Key	Location/Qualifiers
FT	Modified-site 44	/note= "Potential phosphorylation site"
FT	Modified-site 60	/note= "Potential phosphorylation site"
FT	Modified-site 73	/note= "Potential phosphorylation site"
FT	Modified-site 98	/note= "Potential phosphorylation site"
FT	Modified-site 117	/note= "Potential phosphorylation site"
FT	Modified-site 123	/note= "Potential phosphorylation site"
FT	Modified-site 180	/note= "Potential phosphorylation site"
PN	W09964593-A2.	
PD	16-DEC-1999.	
PX		
PX	08-JUN-1999;	99WO-US012906.
PR	08-JUN-1998;	98US-0088695P.
PA	(INCY-) INCYTE PHARM INC.	
PI	Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Baughn MR, Patterson C;	
DR	WPI; 2000-105887/09. N-PSDB; AAZ29482.	
PT	Novel regulatory proteins, for diagnosis, treatment and prevention of cell proliferative and immune system diseases.	
XX		
PS	Claim 1; Page 70; 88pp; English.	
CC	The present sequence is cell cycle regulation protein-3 (CECRP-3). Prints analytical method was used to identify this protein. CECRPs are activators of cell proliferation or inhibitors of cellular processes that modulate proliferation.They are used to treat or prevent cell proliferative diseases like cancers, atherosclerosis, cirrhosis, hepatitis, psoriasis, immune system disorders (e.g. acquired immune deficiency syndrome, allergy, asthma, Crohn's disease, rheumatoid arthritis). Antibodies are raised to screen for specific binding agents. The corresponding nucleic acid is used in gene therapy, chromosomal mapping and isolation of related sequences	
Sequence	236 AA;	
Query Match	20.6%; Score 213; DB 3; Length 236;	
Best Local Similarity	32.6%; Pred. No. 3.2e-12;	
Matches	76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;	
4 GIKRKHSDLSEEE---ERWEWSPALQSYQA-----LIRISLDKVQRSLGPPRS 51		

Db	5	GLKRKREBEKEPLAVDSW-WLDPGHAAVAQAAPAVASSSLFDLSVLKLRHSLQQSEPD	63
QY	52	LRRLVLIHNTLQQLQAALRLAPAPALPEPL-----FLGEEDFSLSATIGSLIRE	101
Db	64	LRHLVLVNTLRRIQAS--MAPAALPVPVSPAPASVADNLLASSDAALASMASLLED	121
QY	102	LDTSMDGTEPPQNPTPLGLQNEVP-----QPDVFLBAL--SSRYLGDSGLDDFFL	152
Db	122	L-SHIEGLSQAPQP-----LADEGEPPGRSIGGAAPSLGALDILGPATGCLDDGLEGLFE	175
QY	153	DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEITLGS	196
Db	176	DIDTSMYDNEIMAPASEGLKPGPED---GPGKEAPELDEAEILDYIMDVIVGT	225

```

RESULT 11
ID AAM93724 standard; protein; 236 AA.
XX AAM93724
AC AAM93724;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 3677.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
DR N-PSDB; AAK94674.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 8; SEQ ID NO 3677; 1380bp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesising the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a polypeptide encoded by a full length
CC human cDNA of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in CD-ROM
CC format directly from EPO
XX
SQ Sequence 236 AA;

```

[illegible]

```

RESULT 12
AAG89292
ID      AAG89292 standard; protein; 236 AA.
XX
XX      AAG89292;
AC
XX      11-SEP-2001 (first entry)
XX      DT
XX      Human secreted protein, SEQ ID NO: 412.
XX      DE
XX      Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW      GENSET.
XX
XX      Homo sapiens.
OS
XX      WO200142451-A2.
XX      PN
XX      14-JUN-2001.
XX      PD
XX      07-DEC-2000; 2000WO-IB001938.
XX      PF
XX      08-DEC-1999; 99US-0169629P.
PR      06-MAR-2000; 2000US-0187470P.
XX
XX      (GEST ) GENSET.
PA
XX      Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX      PI
XX      WPI, 2001-367870/38.
XX      DR
XX      N-PSDB; AAH64895.
XX
XX      Full length GENSET human nucleic acids encoding potentially secreted
PT      proteins, useful in gene therapy and vaccination against a variety of
PT      diseases, and for diagnosis of those diseases.
PT
XX
XX      Claim 21; Page 889-890; 921pp; English.
XX      PS
XX
XX      The invention relates to full length GENSET human nucleic acids encoding
CC      potentially secreted proteins. The nucleic acids and the polypeptides
CC      they encode may be used in the prevention, treatment and diagnosis of
CC      diseases associated with inappropriate GENSET gene expression. For
CC      example, they be used to treat disorders associated with decreased GENSET
CC      gene expression by rectifying mutations or deletions in a patient's
CC      genome that affect the activity of GENSET or by supplementing the
CC      patients own production of GENSET polypeptides. Conversely, antisense
CC      nucleic acid molecules may be administered to down regulate GENSET
CC      expression by binding with the cells' own genes and preventing their
CC      expression. The sense and antisense nucleic acids may also be used as DNA
CC      probes in diagnostic assays to detect and quantitate the presence of
CC      similar nucleic acid sequences in samples, and hence to determine which
CC      patients may be in need of restorative therapy. The GENSET polypeptides
CC      may be used as antigens in the production of antibodies and in assays to
CC      identify modulators (agonists and antagonists) of GENSET polypeptide
CC      expression and activity. The present sequence is a GENSET polypeptide of
CC      the invention
XX
XX      Sequence 236 AA;
SQ

```


	Query Match	20.6%;	Score 213;	DB 4;	Length 236;
	Best Local Similarity	32.6%;	Pred. No. 3.2e-12;		
	Matches	76;	Conservative 36;	Mismatches 69;	Indels 52; Gaps 13;
QY	4 GLKRKHSDLSEEE---ERWENSPAGLQSYYQA-----LRISLDKVQRSLGPPRAPS	51			
Db	5 GLKRKRSEEBEKEPLAVDSW-WIDPGHTAVAQAAPPAVASSSLDPSVLKLHSHLQQSEPD	63			
QY	52 LRRHVLIHNTLQOLQAALRLAPALPEPL-----FLGEEDFSLSATIGSILRE	101			
Db	64 LRHLVLVVNTRLRIQAS--MAPAALPFVPSPAPASVADNLTLSSDAALSASMASLTLED	121			
QY	102 LDTSMDGTBPQNPTPLGLONEVPP-----QPDPVFLEAL--SSRYLGDSGLDDFFL	152			
Db	122 L-SHIRGLSQAPQP-----LADEGPGRSIGGAAPSLGALDLLGPATGCLLDDGLEGLFE	175			
QY	153 DIDTSAVEKE---PARAPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS	196			
Db	176 DIDTSMYDNELMARPASEGLKPGEED--GPGKEAPBLDEAEILDYLMADVLVGT	225			

```

RESULT 13
ADL31644
ID ADL31644 standard; protein; 236 AA.
XX
AC ADL31644;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human protein encoded by a full length cDNA clone SegID 3677.
XX
KW human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method.
XX
OS Homo sapiens.
XX
PN EPI396543-A2.
XX
PD 10-MAR-2004.
XX
PF 07-JUL-2000; 2003EP-00025638.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Negai K, Kojima S, Otsuki T, Koga H;
PI
XX
DR WPI; 2004-204755/20.
DR N-PSDB; ADL31643.
XX
PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
PS
PS Example 1; SEQ ID NO 3677; 1340pp; English.
XX
XX This invention relates to a novel primers useful for synthesizing full
XX length cDNA molecules that encode human proteins. Specifically, it refers
XX to secretory or membrane proteins that are potential therapeutic agents/
XX target molecules in the field of medicine, and in particular genes
XX encoding proteins that are associated with signal transduction,
XX glycoproteins and transcription. The present invention describes a method
XX for efficiently cloning a full length human cDNA from both the 5' and 3'
XX ends using the oligo-capping method. This polypeptide sequence is a full
XX length human protein of the invention.
XX
XX Sequence 236 AA;
XX

```

```

Query Match          20.6%; Score 213; DB 8; Length 236;
Best Local Similarity 32.6%; Pred. No. 3.2e-12;
Matches      76; Conservative    36; Mismatches   69; Indels    52; Gaps     13,
QY      4 GLKRKHSDLSEEE---ERNWSPAGLQSYQA-----LLRISLDKVQRSLGPRAPS 51
        ||||| : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      5 GLKRKREEEKEPLAVDSW-WLDPGHTAVAQAAPPAAVASSSIFDLSVLKLHSLQQSEPD 63
QY      52 LRRHVLIHNTLQQLQAALRLAPAPALPREPL-----FLGEEDFSLSATIGSILRE 101
        || : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      64 LRHLVLVVNTRLRIQAS--MAPAALPVPSPRAPASVADNLIASDAALSASMSTLED 121
QY      102 LDTSMDGTBPQNPFVTPLGLONEVPP-----QPDPVFLEAL--SSRYLGDSGLDFFFL 152
        | : : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      122 L-SHIEGLSQAPQF-----LADEGPGRSIGGAAPSLGALDILGPATGCLLDDGLBGLFE 175
QY      153 DIDTSAVEKE---PARAPPF-PHNLFCAPSWE-----WNELDHIMEIILGS 196
        ||||| : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      176 DDTSMYDNELMAFASIEGLKPQPED---GPKKEAPELDEAEILDYLMDVLVGTT 225

```

RESULT	14
AAM25550	ID AAM25550 standard; protein, 278 AA.
XX AC	AAM25550;
XX DT	16-OCT-2001 (first entry)
XX DE	Human protein sequence SEQ ID NO:1065.
XX KW	Human; cancer; ulcer; HIV infection; human immunodeficiency virus; antiinflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac; central nervous system; virucide;
KW KM	anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia; antiaggregant; haemostatic; vulnery; antilucer; osteopathic; eczema; dermatological; antiallergic; asthma; antidiabetic; cytostatic;
KW KM	neuroprotective; antidepressant; nootropic; antiparkinsonian; infection; immunostimulant; gene therapy; antisense therapy; vaccine; inflammation; antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW KM	cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity; genetic disease; haematopoietic disorder; platelet disorder; asthma; thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW KM	allergic rhinitis; diabetes; multiple sclerosis; depression; Alzheimer's disease; Parkinson's disease; neurodegenerative disorder; neurological disorder.
KW KM	Homo sapiens.
OS XX	WO200153455-A2.
PN XX	26-JUL-2001.
PD XX	22-DEC-2000; 2000WO-US035017.
PF XX	23-DEC-1999; 99US-00471275.
PR XX	21-JAN-2000; 2000US-00488725.
PR PR	25-APR-2000; 2000US-00552317.
XX PA	(HYSE-) HYSEQ INC.
XX PI	Tang YT, Liu C, Drmanac RT;
XX XX	WPI; 2001-457603/49.
DR DR	N-PSDB; AAH99491.
XX PT	Isolated human polynucleotides encoding polypeptides, useful for the treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX PS	Claim 20; Page 214; 1217pp; English.
CC CC	AAH99166 to AAH99904 encode the human proteins given in AAM25225 to AAM25963. The proteins can have activities based on the tissues and cells

CC they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
CC antitumor; osteopathic; dermatological; antiallergic; antiasthmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders

XX Sequence 278 AA;

Query Match 20.6%; Score 213; DB 4; Length 278;
Best Local Similarity 32.6%; Pred. No. 3.9e-12;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;

QY 4 GLKRKHSLEEBE---ERWESPAGLSQSYQA-----LIRISLDKVQRSIGPPAPS 51
DB 47 GLKRKREBEEKEPLAVDSW-WLDPGHAAVAQAAPPAVASSSLFDLSVKLHSHLQOSEPD 105
QY 52 LRRHVLHNTLQQLQAALRLAPALPPEPL-----FLGEBDFSLSATIGSILRE 101
DB 106 LRHLVLVNTLRRIGAS--MAPAALPVPVSPPAAPSVADNLLASSDAALSASMASLLED 163
QY 102 LDTSMGTBPPQNPVTPFLGLQNEVPP-----QPPPVFLAAL--SSRYLGDSGLDDFFL 152
DB 164 L-SHIEGLSQAPQF-----LADGPPGRSIGGAAPSLGALDILGPATGCLDDGLEGLFE 217
QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196
DB 218 DIDTSMYDNLWAPASEGLKPGPED---GPGKEAPELDEAELDYLMDVLVGT 267

RESULT 15

AAV31829 ID AAV31829 standard; protein; 236 AA.

XX AC AAV31829;

XX DT 06-DEC-1999 (first entry)

XX DE Human adult blood secreted protein g21_1.

XX KM Secreted protein; g21_1; human; therapy; diagnosis; vaccine; blood.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT FT Misc-difference 1 /note= "encoded by AWG"

FT FT Peptide 19..21 /note= "signal peptide"

FT FT Protein 32..236 /note= "mature protein"

FT FT Domain 40 /note= "a putative transmembrane domain is centered around this residue"

FT FT Peptide 67..79 /note= "alternative signal peptide"

FT FT Protein 80..236 /note= "alternative mature protein"

FT FT Domain 80 /note= "a putative transmembrane domain is centered

FT around this residue"
FT Misc-difference 137 /note= "encoded by GAS"
FT Domain 150 /note= "a putative transmembrane domain is centered around this residue"

XX PN W09947555-A1.

XX PD 23-SEP-1999.

XX PF 18-MAR-1999; 99WO-US005939.

XX PR 20-MAR-1998; 98US-0078803P.

XX PR 17-MAR-1999; 99US-00078803.

XX PA (GENY) GENETICS INST INC.

XX PI Jacobs K, Mccoy JM, Lavalie ER, Collins-Racie LA, Evans C;

XX PI Merberg D, Treacy M, Agostino MJ, Steindinger RJ;

XX DR WPI; 1999-562059/47.

XX DR N-PSDB; AAZ19894.

PT New polynucleotides derived from murine fetal cell cDNA libraries,

PS potentially used as, e.g. vaccines.

PS Claim 13(a); Page 94; 107pp; English.

CC This is the predicted amino acid sequence of a novel human secreted
CC protein, g21_1, as deduced from an isolated adult blood cDNA clone (see
CC AAZ19894). The invention provides new human secreted proteins (see
CC AAV31828-38) and polynucleotides (see AAZ1993-901) isolated from foetal
CC cell, adult blood, adult brain and foetal kidney cDNA libraries. They are
CC predicted to have biological activities which would make them suitable
CC for treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data are given. Suggested activities
CC include nutritional, cytokine, tissue growth, cell proliferation and
CC differentiation, immunostimulant (e.g. as vaccine), immunosuppressive,
CC haematopoiesis regulating, activin or inhibin, chemotactic or
CC chemokinetic, haemostatic or thrombolytic, receptor/ligand activity,
CC antiinflammatory, cadherin or tumour invasion suppressor, and tumour
CC inhibition activities

XX SQ Sequence 236 AA;

Query Match 20.4%; Score 211; DB 2; Length 236;
Best Local Similarity 32.6%; Pred. No. 5e-12;
Matches 76; Conservative 35; Mismatches 70; Indels 52; Gaps 13;

QY 4 GLKRKHSLEEBE---ERWESPAGLSQSYQA-----LIRISLDKVQRSIGPPAPS 51

DB 5 GLKRKREBEEKEPLAVDSW-WLDPGHAAVAQAAPPAVASSSLFDLSVKLHSHLQOSEPD 63

QY 52 LRRHVLHNTLQQLQAALRLAPALPPEPL-----FLGEBDFSLSATIGSILRE 101

DB 64 LRHLVLVNTLRRIGAS--MAPAALPVPVSPPAAPSVADNLLASSDAALSASMASLLED 121

QY 102 LDTSMGTBPPQNPVTPFLGLQNEVPP-----QPPPVFLAAL--SSRYLGDSGLDDFFL 152

DB 122 L-SHIEGLSQAPQF-----LAXEGPPGRSIGGAAPSLGALDILGPATGCLDDGLEGLFE 175

QY 153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196

DB 176 DIDTSMYDNLWAPASEGLKPGPED---GPGKEAPELDEAELDYLMDVLVGT 225

Search completed: February 6, 2006, 09:56:50
Job time : 203 secs

This Page Blank (uspto)

Db 634 ASLHLDVLPSSNLEKPPSPVAAPPLPTFSAPSLPQGSVSTSI P SPPV-APTLSTR-- 690
Qy 143 GDSGLDDFFLDIDTSAVEKEPARAPPEPP 171
Db 691 -----TETESISKNPTKSPPEPP 708

RESULT 5

A42029
transcription factor B3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A42029
R;Roman, C.; Macera, A.G.; Cooper, C.; Artandi, S.; Blain, S.; Ward, D.C.; Calame, K.
Mol. Cell. Biol. 12, 817-827, 1992
A;Title: mTFE3, an X-linked transcriptional activator containing basic helix-loop-helix
A;Reference number: A42029; MUID:92123207; PMID:1732746
A;Accession: A42029
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-446 <ROM>
A;Cross-references: UNIPROT:Q64092; UNIPARC:UPI00000296F2; GB:S76673; NID:g243439; PIDN:
A;Note: sequence extracted from NCBI backbone (NCBIN:76673, NCBI P:76674)

Query Match 9.2%; Score 95.5; DB 2; Length 446;
Best Local Similarity 27.3%; Pred. No. 1.7;
Matches 48; Conservative 24; Mismatches 61; Indels 43; Gaps 8;

Qy 12 LEEERERWESPAGLQSYQOALRLSLDKVQSLGPRAPSLRRHVLHNTLQQLQALRL 71
Db 272 LQKQQR---SKLESRRQ-----SLEQANSLQLRIQL-----ELQALHIG 311
Qy 72 APAPALPEPLFLGEEDFSLSATIGSLRELDTSMGTEPPQNPVTPLGLQNEVPPQEDP 131
Db 312 LP---VPPNPGLSLTSSVSDSLKP--EQLDIEEGRPSTTFHVSGGPAQANAPQQPPA 366
Qy 132 VFLEAL-----SSRYLGDSGLDDFFLDIDTSAVEKE-----PARAPPEP 170
Db 367 PRSDALLDLHPPSDHLGLDG-DPFLGLELDIMEEGMVGSLGALSPLRAASDP 421

RESULT 6

A54602
microtubule-associated serine/threonine protein kinase MAST205 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Dec-2004
C;Accession: A54602
R;Walden, P.D.; Cowan, N.J.
Mol. Cell. Biol. 13, 7625-7635, 1993
A;Title: A novel 205-kilodalton testis-specific serine/threonine protein kinase associat
A;Reference number: A54602; MUID:94067123; PMID:8246979
A;Accession: A54602
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1734 <WAL>
A;Cross-references: UNIPROT:Q60592; UNIPARC:UPI0000027F86; GB:U02313; NID:g406057; PIDN:
C;Keywords: ATP
F;451-726/Domain: protein kinase homology <KIN>
F;459-467/Region: protein kinase ATP-binding motif

Query Match 9.1%; Score 93.5; DB 2; Length 1734;
Best Local Similarity 26.3%; Pred. No. 15;
Matches 51; Conservative 23; Mismatches 71; Indels 49; Gaps 8;

Qy 25 GLQSYQOAL-----LRISLDKVQSLGPRAPSLRRHVLHNTLQQLQALRLAPAPALPP 79
Db 1308 GSQGFPTKLHLSPPLGRQLSRKSAEPPRSPILKR-----VQSAEKLAALAAAE 1357
Qy 80 EPLFLGEEDFSLSATIGSLRELDTSMGTEPPQNPVTPLGLQNEVP---PQDPVFLEA 136
Db 1358 KKL-APSRKHSLDLPHGELKKEI-----TPREASPLVAVGTRSVLSGKGPLPGKGVLP 1410

Qy 137 LSSRYLG-----DSGLDDFFLDIDTSAVEKEPARAP-PEPPHN 173
Db 1411 APSRALGTLRQDRAERRESLQKQAIKRVDSSEDDTDEEPENSQATQEPRLSPHPASHN 1470
Qy 174 LFCAPGSWEWNELD 187
Db 1471 LL-PKSGEGTEED 1483

RESULT 7

FOLJH2
gag polyprotein - human T-cell lymphotropic virus type 2
N;Alternate names: core polyprotein
N;Contains: core protein p12; core protein p15; core protein p24
C;Species: human T-cell lymphotropic virus type 2, HTLV-2
A;Note: host Homo sapiens (man)
C;Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 09-Jul-2004
C;Accession: A03944
R;Shimotohno, K.; Takahashi, Y.; Shimizu, N.; Gojobori, T.; Golde, D.W.; Chen, I.S.Y.; I
Proc. Natl. Acad. Sci. U.S.A. 82, 3101-3105, 1985
A;Title: Complete nucleotide sequence of an infectious clone of human T-cell leukemia vi
A;Reference number: A94042; MUID:85216449; PMID:2582407
A;Accession: A03944
A;Molecule type: DNA
A;Residues: 1-433 <SHI>
A;Cross-references: UNIPROT:P03346; UNIPARC:UPI000012B002; GB:M10060; NID:g329559; PIDN
A;Note: the authors translated the codon TAC for residues 197 and 249 as Thr
C;Genetics:
A;Gene: gag
C;Superfamily: mammalian retrovirus gag polyprotein II
C;Keywords: core protein, polyprotein
F;1-136/Product: core protein p15 #status predicted <P15>
F;137-214/Product: core protein p24 #status predicted <P24>
F;215-433/Product: core protein p12 #status predicted <P12>

Query Match 9.0%; Score 93; DB 1; Length 433;
Best Local Similarity 23.1%; Pred. No. 2.7;
Matches 40; Conservative 24; Mismatches 67; Indels 42; Gaps 7;

Qy 47 PRAPSLRRHVLHNTLQQLQALRLAPAP-----LPPEPLFLGEEDFSLSA 93
Db 13 PKAP--RGLSTHMLNPLQAAVRLQPRPSDFDQQLRRFLKALKTPIMWLPIDYSLIA 69
Qy 94 TI-----GSILRELDTSMGTEPPQNPVTPLG---LQNEVPPQDPVFLEA-LSSRYLG 143
Db 70 SLIPKGYPRVVEIINLVKNQVSPAPAPVPTPICPTTTPPPPPSPKAAHVPPY-- 127
Qy 144 DSGLDDEFLDIDTSAVEKEPARAPPEPPHNLFCAPGSWEWNELDHIMEIILGS 196
Db 128 -----VEPTTQCPILHPPGAP-----SAHRRPMQKDLQAIKQEVSSS 166

RESULT 8

A82255
hypothetical protein VC0998 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82255
R;Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: A82255
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1621 <HBI>
A;Cross-references: UNIPROT:Q9KTA5; UNIPARC:UPI00000C2B4C; GB:AB004181; GB:AB003852; NID:
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0998
A;Map position: 1

A;Residues: 1-1676 <FIS>
A;Cross-references: UNIPROT:Q00083; UNIPARC:UP10000125CBB; GB:X82289; MID:g683499; PID:5
C;Genetics:
A;Gene: *apsA*
A;Introns: 149/3

Query Match	8.5%	Score 88;	DB 2;	Length 1676;
Best Local Similarity	20.5%	Pred. No. 41;		
Matches 32;	Conservative 19;	Mismatches 51;	Indels 54;	Gaps 6;

```

QY      36  ISLDKVQKSLGPRAPSLRRHVLINHTLQQLQALRLAPALPPEPLFLGGEEDSLSATI  95
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      837  VSFEETE-PVAPSFPELRTAFVGVGCTTEPVAAAPVVPPEVALSP-----I  880

QY      96  GSILRELDTSMDGTEPPQNPTPLGLQNEVPQDPVFL-----  134
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      881  SS-----QTTQPTPEVIP-----APPEPEPIYVPEMAFSQLVEDTFLILAKLP  924

QY      135  EALSSRYLGDSGLDDFFLDIDITSAVEKEPARAPPEP  170
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      925  EPAPERVFVFAEQGTSTDIADVSAISSEQTE-PVEP  959

```

RESULT 13
A34596
transcription factor E3 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Jul-1990 #sequence_revision 09-Oct-1992 #text_change 09-Jul-2004
C;Accession: A34596; S10379
R;Beckmann, H.; Su, L.K.; Kadesch, T.
Genes Dev. 4, 167-179, 1990
A;Title: TFEB: a helix-loop-helix protein that activates transcription through the immunophilin binding site
A;Reference number: A34596; MUID:90249724; PMID:2338243
A;Accession: A34596
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-536 <BEC>
A;Cross-references: UNIPROT:P19532; UNIPARC:UPI000016B0B0; EMBL:X51330; NID:g37061; PIDD:A34596
A;Note: the authors translated the codon ACC for residue 433 as Ser, and GAG for residue 434 as Glu
C;Genetics:
A;Gene: GDB:TFEB3
A;Cross-references: GDB:125870; OMIM:314310
A;Map position: Xp11.23-Xp11.22
C;Keywords: DNA binding; transcription factor

Query Match	8.4%	Score 86.5;	DB 2;	Length 536;
Best Local Similarity	25.4%;	Pred. No. 13;		
Matches	47;	Conservative	24;	Mismatches 55; Indels 59; Gaps 10;
QY	12	LEEEERWESPAGLQSYQOALRISLDKVRQSLGPRAPSLRRHVLIHNTLQOLQALRL	71	
		:::	::	::
Db	192	LQEQQR---SKLRSRQR---SLEQANRSLQLRIGEL-----ELQQAQIHG	231	
QY	72	APAPALPREPLFLGEBDFSLSATIGSILRE-LDTSMDGTE-----PPONPVTPLG	120	
Db	232	LPVPGTP-----GLSLATSTSTDSLKPEQLDIEEBGRPGARTFHVGGGPAQN-----	279	
QY	121	LQNEVFPQP--DPVFLEALSSRYLGDGSLDDFFLDIDTSAVEKE-----PAR	165	
		::	:	::
Db	280	APHQOPPARPPSDALLDHFPSDHLGDLG-DPFLGLELDILMEEEGVVGGLSGALSPLR	338	
QY	166	APPEP	170	
			:	
Db	339	AASDP	343	

RESULT 14
T14355
protein-tyrosine-phosphatase (EC 3.1.3.48) TD14 - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: T14355
R/Cao, L.; Zhang, L.; Ruiz-Lozano, P.; Yang, Q.; Chien, K.R.; Graham, R.M.; Zhou, M.

J. Biol. Chem. 273, 21077-21083, 1998
 A;Title: A novel putative protein-tyrosine phosphatase contains a BROI-like domain and s
 A;Reference number: Z18004; MUID:98361981; PMID:9694860
 A;Accession: T14355
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-1494 <CAO>
 A;Cross-references: UNIPROT:O88902; UNIPARC:UPI00000E5DAD; EMBL:AF077000; NID:g3598973;
 A;Experimental source: brain
 C;Genetics:
 A;Gene: PTP-TD14
 C;Function:
 A;Description: may be involved in regulating Ha-ras-dependent cell growth
 C;Keywords: phosphoric monoester hydrolase

	Query Match	8.4%;	Score 86.5;	DB 2;	Length 1494;	
	Best Local Similarity	29.1%;	Pred. No. 47;			
	Matches	37;	Conservative	19;	Mismatches	40; Indels 31; Gaps 9;
QY	62 LQOLQALRLAPADALPEEPFLGEBEDFSLATGSILRE-----LDTSMGDTEPPQN	P	115			
Dd	625 LMQPRAAVPMAPGPVLYPAFVTSE-----LGIVPRSSPQHGISSPYAGVGPPQ-P	P	675			
QY	116 VTPIGLQNEVPQQ---PDVFLEALSSRYLGDGDLDDFFLDIDT-SAVEKEPARAPEPP	P	171			
Dd	676 I--VGLPSAPPQFSGR-----LAMDV-R-PATTVDVSQAIPISSHMALRPGAPAPPPQ-	P	727			
QY	172 HNLFCAP	178				
Dd	728 ----CFP	730				

RESULT 15
B37761
ntrA protein - Thiobacillus ferrooxidans
C;Species: Thiobacillus ferrooxidans
C;Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 09-Jul-2004
C;Accession: B37761
R;Berger, D.K.; Woods, D.R.; Rawlings, D.E.
J. Bacteriol. 172, 4399-4406, 1990
A;Title: Complementation of Escherichia coli sigma(54) (ntrA)-dependent formate hydrogen
A;Reference number: A37761; MUID:90330545; PMID:2198257
A;Accession: B37761
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-475 <BER>
A;Cross-references: UNIPROT:P24695; UNIPARC:UPI0000134582; GB:M58480; GB:M33831; NID:915
C;Superfamily: Pseudomonas transcription initiation factor sigma
C;Keywords: DNA binding; transcription regulation

```

Query Match      8.3%; Score 86; DB 2; Length 475;
Best Local Similarity 21.7%; Pred. No. 12;
Matches 40; Conservative 31; Mismatches 75; Indels 38; Gaps 6;

QY      1 MEGGLKRKHSDL EEEERWESPAGLSYQQALRLISLDKVOR----SLGPRA--PSLR 53
          ||| | : :: | | | | | : : : | : | : | : | :
Db      307 MAGGKDAHKYIDQLNEARWFIKSLQSRODTILKVARAIVEROKDFANGPESMRPVTL 366

QY      54 RHVLINHTLQQLQAALRLAPALRPEPLFLGEEDSLSATIG-----SILREIDT 104
          ||| | : : : : | | | | | : : : | : | : | : | :
Db      367 RH--IADAVEMHSESTVSRVTNQKMITPRGLYEKFYFSSHVGTDSGSASATAIRALLI 424

QY      105 SMDGTEPQNVPVTPGLQNEVPPQPDPVFLEALSRRYLGDGSLDDFFLLDIDTSAVEKEPA 164
           | | | : : : : : : : | : | | | : : | : | : | :
Db      425 KMTQABDAQHPLS-----DATAIRVLADGGIQ--IARRTVAKYREAA 464

QY      165 RAPP 168
           ||
Db      465 NVPP 468
```

Search completed: February 6, 2006, 10:01:57
Job time : 46 secs

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 09:53:36 ; Search time 253 Seconds
(without alignments)
546.575 Million cell updates/sec

Title: US-10-069-386A-2
Perfect score: 1033
Sequence: 1 MEGGLKRXHSDLEEEBERNE.....APGSWEWNELDHIMEILGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	1026	99.3	196	1 SRTD3_HUMAN	Q9ujw9 homo sapien
2	871	84.3	197	1 SRTD3_MOUSE	Q9erc3 mus musculu
3	858	83.1	197	2 Q5BK27_RAT	Q5bk27 rattus norv
4	217	21.0	236	1 SRTD1_MOUSE	Q9j110 mus musculu
5	216.5	21.0	244	2 Q58EQ0_BRARE	Q58eq0 brachydanio
6	213	20.6	236	1 SRTD1_HUMAN	Q9uhv2 homo sapien
7	213	20.6	236	2 Q53GC0_HUMAN	Q53gc0 homo sapien
8	208	20.1	236	2 Q6P771_RAT	Q6p771 rattus norv
9	166	16.1	244	2 Q9DCZ2_MOUSE	Q9dcz2 mus musculu
10	163.5	15.8	244	2 Q4T2W4_TETNG	Q4t2w4 tetraodon n
11	145	14.0	275	2 Q4S6F1_TETNG	Q4s6f1 tetraodon n
12	144.5	14.0	229	2 Q5RIK3_BRARE	Q5rik3 brachydanio
13	140.5	13.6	309	1 SRTD2_MOUSE	Q9jfg5 mus musculu
14	140.5	13.6	309	2 Q5SSE5_MOUSE	Q5sse5 mus musculu
15	136.5	13.2	248	2 Q6GM81_XENLA	Q6gm81 xenopus lae
16	136	13.2	237	1 CDCA4_MOUSE	Q9cwm2 mus musculu
17	134.5	13.0	315	2 Q4V7A7_RAT	Q4v7a7 rattus norv
18	133	12.9	361	2 Q6NXP9_BRARE	Q6nxd9 brachydanio
19	133	12.9	383	2 Q7ZZ27_BRARE	Q7zz27 brachydanio
20	132	12.6	316	2 Q5ZUE5_CHICK	Q5zje5 gallus gall
21	130.5	12.6	286	2 Q5RIK2_BRARE	Q5rik2 brachydanio
22	125	12.1	314	1 SRTD2_HUMAN	Q14140 homo sapien
23	125	12.1	314	2 Q53TS2_HUMAN	Q53ts2 homo sapien
24	125	12.1	314	2 Q5RDH4_PONPY	Q5rhd4 pongo pygma
25	121	11.7	364	2 Q4SG65_TETNG	Q4sg65 rattus norv
26	115	11.1	1194	2 Q76M68_RAT	Q76m68 rattus norv
27	113.5	11.0	911	2 Q8OTJ8_MOUSE	Q8otj8 mus musculu
28	113	10.9	249	2 Q5ZL47_CHICK	Q5zl47 gallus gall
29	112.5	10.9	246	2 Q6DFI1_XENLA	Q6dfi1 xenopus lae
30	105.5	10.2	757	1 CEP68_HUMAN	Q76n32 homo sapien
31	105	10.2	241	1 CDCA4_HUMAN	Q9bx18 homo sapien

32	104.5	10.1	740	2 Q9UPB2_HUMAN	Q9upp2 homo sapien
33	101.5	9.8	728	1 P85B_HUMAN	Q00459 homo sapien
34	101.5	9.8	728	1 Q5EAT5_HUMAN	Q5eat5 homo sapien
35	100.5	9.7	361	2 Q5STY1_CRYNE	Q5sty1 cryptococcu
36	100.5	9.7	361	2 Q5KIF7_CRYNE	Q5kif7 cryptococcu
37	100	9.7	846	2 Q6NVC9_BRARE	Q6nvc9 brachydanio
38	99	9.6	798	2 Q571B6_MOUSE	Q571b6 mus musculu
39	98.5	9.5	314	1 SOX12_MOUSE	Q04890 mus musculu
40	98	9.5	278	2 Q55XN0_CRYNE	Q55xn0 cryptococcu
41	98	9.5	517	2 Q9EEF8_NPVAG	Q9eef8 anticarsia
42	98	9.5	562	2 Q8NAF0_HUMAN	Q8naf0 homo sapien
43	98	9.5	768	2 Q5BEH7_EMENTI	Q5beh7 aspergillus
44	98	9.5	1865	2 Q5VU37_HUMAN	Q5vu37 homo sapien
45	98	9.5	1884	2 Q4RWK6_TETNG	Q4rwk6 tetraodon n

ALIGNMENTS

RESULT 1
SRTD3_HUMAN STANDARD; PRT; 196 AA.
ID SRTD3_HUMAN AC Q9UJW9; Q96CQ2;
DT 01-FEB-2005 (Rel. 46, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE SERTA domain-containing protein 3 (Replication protein-binding trans-activator) (RPA-binding trans-activator).
GN Name=SBRTAD3; Synonyms=RBTL1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, AND INTERACTION WITH RPA2.
RX MEDLINE=20440390; PubMed=10982866; DOI=10.1093/nar/28.18.3478; Cho J.M., Song D.J., Bergeron J., Benlilame N., Wold M.S., Alaoui-Jamali M.A.;
RA "RBTL1, a novel transcriptional co-activator, binds the second subunit of replication protein A.";
RT Nucleic Acids Res. 28:3478-3485(2000).
RL
RN
RP
RC TISSUE=Ovary, and Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899; Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Shevchenko Y., Bouffard G.G., Whiting M., Madan A., Young A.C., Green E.D., Dickson M.C., Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.B., Scherch A., Schein J.E., Jones S.J.M., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Strong transcriptional co-activator.
CC -!- SUBUNIT: Interacts with RPA2.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Contains 1 SERTA domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC -----
DR EMBL; AF192529; AAP05761.1; -; mRNA.
DR EMBL; BC014061; AAH14061.1; -; mRNA.
DR EMBL; BC050643; AAH50643.1; -; mRNA.
DR Ensembl; ENSG00000167565; Homo sapiens.
DR HGNC; HGNC:17931; SERTAD3.
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA; 1.
DR PROSITE; PS51053; SERTA; 1.
KW Activator; Nuclear protein; Transcription; Transcription regulation.
FT DOMAIN 26 73 SERTA.
FT CONFLICT 2 2 V -> E (in Ref. 1).
SQ SEQUENCE 196 AA; 21769 MW; 9C54AFA126F17AF1 CRC64;
Query Match 99.3%; Score 1026; DB 1; Length 196;
Best Local Similarity 99.5%; Pred. No. 1.3e-70;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MEGGLKRXHSDLEBBERRWSPAGLQSYQALRLISLDKVQSLGPRAPSLRRHYLIHN 60
DB 1 MVGGLKRXHSDLEBBERRWSPAGLQSYQALRLISLDKVQSLGPRAPSLRRHYLIHN 60
OY 61 TLQQLQAALRLAPAPALPPPLFLGEEDFSLSATIGSLIRELDTSMGTEPPQNPVTPLG 120
DB 61 TLQQLQAALRLAPAPALPPPLFLGEEDFSLSATIGSLIRELDTSMGTEPPQNPVTPLG 120
OY 121 LQNEVPPQDPVFLBALSRYLGDSDLDFLIDITSAVEKEPARAPPEPHNLFCAFGS 180
DB 121 LQNEVPPQDPVFLBALSRYLGDSDLDFLIDITSAVEKEPARAPPEPHNLFCAFGS 180
OY 181 WEWNELDHIMEIILGS 196
DB 181 WEWNELDHIMEIILGS 196

RESULT 2
SRTD3_MOUSE
ID SRTD3_MOUSE STANDARD; PRT; 197 AA.
AC O9ERC3;
DT 01-FEB-2005 (Rel. 46, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE SERTA domain-containing protein 3 (Replication protein-binding trans-
DE activator) (RPA-binding trans-activator).
GN Name=Sertad3; Synonyms=Rbt1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Cho J.M., Alaoui-Jamali M.A.;
RT *Cloning of mouse RBT1 cDNA.*;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maira M.A.,
RT *Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.*;
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Strong transcriptional co-activator (By similarity).
CC -1- SUBUNIT: Interacts with RPA2 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Contains 1 SERTA domain.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC -----
DR EMBL; AF317202; AAG30951.1; -; mRNA.
DR EMBL; BC034886; AAH34886.1; -; mRNA.
DR Ensembl; ENSMUSG0000055200; Mus musculus.
DR MGI; MGI:2180697; Sertad3.
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA; 1.
DR PROSITE; PS51053; SERTA; 1.
KW Activator; Nuclear protein; Transcription; Transcription regulation.
FT DOMAIN 27 74 SERTA.
FT SEQUENCE 197 AA; 21935 MW; CFF6FA4C0A91D7E7 CRC64;
SQ SEQUENCE 197 AA; 21935 MW; CFF6FA4C0A91D7E7 CRC64;

Query Match 84.3%; Score 871; DB 1; Length 197;
Best Local Similarity 85.7%; Pred. No. 9e-59;
Matches 168; Conservative 8; Mismatches 18; Indels 2; Gaps 1;

OY 3 GGLKRXHSDL--BBERRWSPAGLQSYQALRLISLDKVQSLGPRAPSLRRHYLIHN 60
DB 2 GGLKRXHSDLBBERRWSPAGLQSYQALRLISLDKVQSLGPRAPSLRRHYLIHN 61
OY 61 TLQQLQAALRLAPAPALPPPLFLGEEDFSLSATIGSLIRELDTSMGTEPPQNPVTPLG 120
DB 62 TLQQLQAALRLAPAPALPPPLFLGEEDFSLSATIGSLIRELDTSMDEMEPLNPAAGSS 121
OY 121 LQNEVPPQDPVFLBALSRYLGDSDLDFLIDITSAVEKEPARAPPEPHNLFCAFGS 180
DB 122 PQNEIVSQADPVFLBALSRYLGDSDLDFLIDITSAVEKDVALLPPPPHSLFCFGS 181
OY 181 WEWNELDHIMEIILGS 196
DB 182 WEWNELDHIMEIILGS 197

RESULT 3
Q5BK27_RAT
ID Q5BK27_RAT PRELIMINARY; PRT; 197 AA.
AC Q5BK27;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DE Hypothetical protein MGCI08974.
GN Name=MGCI08974;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Ronaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RG NIH MGC Project;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC091229; AAH91229.1; -, mRNA.
DR InterPro; IPR011029; DEATH like.
DR InterPro; IPR003139; D_retro_matrix.
DR InterPro; IPR000524; HTH GntR.
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA; 1.
DR PROSITE; PS51053; SERTA; 1.
KW Hypothetical protein.
SQ SEQUENCE 197 AA; 21924 MW; 5F1586511550A9F2 CRC64;

Query Match 83.1%; Score 858; DB 2; Length 197;
Best Local Similarity 84.2%; Pred. No. 8.9e-58;
Matches 165; Conservative 10; Mismatches 19; Indels 2; Gaps 1;

QY 3 GGLKRGKHSDLBEE--ERWENSPAGLQSYQOALLRISLDKVRSLGPRAPSLRRHYLIHN 60
Db 2 GGLKRGKHSDLBEEDEKMDKMSPTALRSYQOALLRISLDKVRSLGPRAPSLRRHYLIHN 61
QY 61 TLQQLQALRLAPAPALPPPEPLFLGEBDFSLSATIGSILRELDTSMDGTEPPQNPVPLG 120
Db 62 TLQQLQALRLAPAPALPPPEPLFLGEBDFSLSTTIGSILRELDTSDELEPPLNPAAPG 121
QY 121 LQNEVPPQDPVFLFALSSRYLGDGLDFLDIDTSAVEKEPARAPPEPPHNLFCAPGS 180
Db 122 PHDEIMSQADPVFLFALSSRYLGDGLDFLDIDTSAVERDTALPPRPPhSLFCSPGS 181
QY 181 WEWNELDHIMEIILGS 196
Db 182 WEWNELDHIMEIILGS 197

RESULT 4
SRTD1 MOUSE STANDARD; PRT; 236 AA.
AC Q9JUL10; Q925E6; Q9D888;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE SERTA domain-containing protein 1 (Transcriptional regulator
DE interacting with the PHD-bromodomain 1) (TRIP-Br1) (CDK4-binding
DE protein p34SEI1) (SEI-1).
GN Name=Sertadi; Synonyms=Sei1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ohtani N., Hara E.;

RT "Cloning of mouse SEI-1 cDNA."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE, FUNCTION, AND INTERACTIONS.
RX MEDLINE=21231173; PubMed=11331592; DOI=10.1093/emboj/20.9.2273;
RA Hsu S.-I., Yang C.M., Sim K.G., Hentschel D.M., O'Leary E.,
RA Bonventre J.V.;
RT "TRIP-Br: a novel family of PHD zinc finger- and bromodomain-
RT interacting proteins that regulate the transcriptional activity of
RT E2F-1/DP-1.";
RL EMBO J. 20:2273-2285 (2001).

RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Embryo, and Small intestine;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Forrest A., Frazer K.S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Zhu Y., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).

RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Ronaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Acts at E2F-responsive promoters to integrate signals
CC provided by PHD- and/or bromodomain-containing transcription
CC factors. Stimulates E2F-1/DP-1 transcriptional activity. Renders
CC the activity of cyclin D1/CDK4 resistant to the inhibitory effects
CC of p16(INK4a).
CC -!- SUBUNIT: Interacts with the PHD-bromodomain of TIF1, TRIM28/TRIFB

RL Genes Dev. 13:3027-3033(1999).
[2]
RP NUCLEOTIDE SEQUENCE, FUNCTION, INTERACTIONS, AND VARIANT ALA-31.
RX MEDLINE=21231173; PubMed=11331592; DOI=10.1093/emboj/20.9.2273;
RA Hsu S.-I., Yang C.M., Sim K.G., Hentschel D.M., O'Leary E.,
RA Bonventre J.V.;
RT "TRIP-Br: a novel family of PHD zinc finger- and bromodomain-
interacting proteins that regulate the transcriptional activity of
RT E2F-1/DP-1.";
RL EMBO J. 20:2273-2285(2001).
[3]
RP NUCLEOTIDE SEQUENCE (GENOMIC DNA).
RA Rieder M.J., Braun A.C., Montoya M.A., Chung M.-W., Nguyen C.P.,
RA Nguyen D.A., Livingston R.J., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RT "NIHS-SNPs, environmental genome project, NIHS ES15478, Department
of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP TISSUE=Uterus;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[1]
CC FUNCTION: Acts at E2F-responsive promoters to integrate signals
provided by PHD- and/or bromodomain-containing transcription
factors. Stimulates E2F-1/DP-1 transcriptional activity. Renders
the activity of cyclin D1/CDK4 resistant to the inhibitory effects
of p16(INK4a).
CC SUBUNIT: Interacts with the PHD-bromodomain of TIF1, TRIM28/TIF1B
and p300/CBP. Binds to DPL. Also interacts with CDK4.
CC SIMILARITY: Contains 1 SERTA domain.
[11]
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
[12]
CC EMBL; AF117959; AAF08349.1; -; mRNA.
DR EMBL; AF366402; AAK52831.1; -; mRNA.
DR EMBL; AY130860; AAM77800.1; -; Genomic DNA.
DR EMBL; BC002670; AAH02670.1; -; mRNA.
DR Ensemble; ENSG00000197019; Homo sapiens.
DR HGNC; HGNC:17932; SERTA1.
DR H-InvDB; HIX0015136; -.
DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
DR GO; GO:0000079; P:regulation of cyclin dependent protein kina. . .; TAS.
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA; 1.
DR PROSITE; PS51053; SERTA; 1.
KW Polymorphism; Transcription; Transcription regulation.
FT DOMAIN 38 85 SERTA.
FT VARIANT 31 31 T -> A (in dbSNP:268687).
FT /FTID=VAR_015881.

SQ SEQUENCE 236 AA; 24704 MW; 8659505783FBF50B CRC64;
Query Match 20.6%; Score 213; DB 1; Length 236;
Best Local Similarity 32.6%; Pred. No. 2.3e-08;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;
QY 4 GLKRKHSDLLEE---ERWESPAQLQSYQQA-----LIRISLDKVQSLGPRAPS 51
DB 5 GLKRKEEERKEEPLAVDSW-WLDPGHTAVAQAAPAVASSSLFDLSVLKIHSLQOSEPD 63
QY 52 LRRHYLIHNTLQQLQAALRLAPALPPEPL-----FLGEDFSLSATIGSLRE 101
DB 64 LRRHLVYVNTLRRIQAS--MAPAALPFVPSPPAPSVADNLLASSDAALSASMASLLED 121
QY 102 LDTSMGTEPPQNPTPLGLQNEVPE-----QPDVFLEAL--SSRYLGDSGLDFFL 152
DB 122 L-SHTEGLSQARP-----LADEGPPGRSIGAAPSIGALDLGPATGCLLDGLGLFE 175
QY 153 DDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIILGS 196
DB 176 DDTSMYDNEMLWAPASEGLKGPED--GPGKEAPELDEALDYLMVLYGT 225

RESULT 7
Q53GC0 HUMAN
ID Q53GC0_HUMAN PRELIMINARY; PRT; 236 AA.
AC Q53GC0;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE SERTA domain containing 1 variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Human small intestine;
RA Maruyama K., Sugano S.;
RT "Oligo-capping: a simple method to replace the cap structure of
RT eucaryotic mRNAs with oligoribonucleotides.";
RL Gene 138:171-174(1994).
[2]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Human small intestine;
RC TISSUE=Human small intestine;
RA Suzuki Y., Yoshitomo K., Maruyama K., Suyama A., Sugano S.;
RT "Construction and characterization of a full length-enriched and a 5'-
RT end-enriched cDNA library.";
RL Gene 200:149-156(1997).
[3]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Human small intestine;
RC TISSUE=Human small intestine;
RA Suzuki Y., Sugano S., Totoki Y., Toyoda A., Takeda T., Sakaki Y.,
RA Tanaka A., Yokoyama S.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK223011; BAD96731.1; -; mRNA.
DR EMBL; AK222970; BAD96690.1; -; mRNA.
FT NON_TER 1 1
SQ SEQUENCE 236 AA; 24703 MW; 8659505783FBF50B CRC64;
Query Match 20.6%; Score 213; DB 2; Length 236;
Best Local Similarity 32.6%; Pred. No. 2.3e-08;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;
QY 4 GLKRKHSDLLEE---ERWESPAQLQSYQQA-----LIRISLDKVQSLGPRAPS 51
DB 5 GLKRKEEERKEEPLAVDSW-WLDPGHTAVAQAAPAVASSSLFDLSVLKIHSLQOSEPD 63
QY 52 LRRHYLIHNTLQQLQAALRLAPALPPEPL-----FLGEDFSLSATIGSLRE 101
DB 64 LRRHLVYVNTLRRIQAS--MAPAALPFVPSPPAPSVADNLLASSDAALSASMASLLED 121

```

QY      102 LDTSMDCGTEPPONPVTPPLGLONEVPP-----QPDPVFLEAL--SSRYLGDSGLDDFFL 152
      | : : : | : : : | : : : | : : : |
Db      122 L-SHIEGLSQAPQP-----LADGCPGRSIGGAAPSLGALDLGPATGCLLDGLEGLEF 175
      | : : : | : : : | : : : | : : : |
QY      153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEILIGS 196
      | : : : | : : : | : : : | : : : |
Db      176 DIDTSMYDNEIWPASEGLKCPGED---GPKKEAPBLDEABLDYLMADVLT 225

```

```

RESULT 8
O6P771 RAT
ID Q6P771_RAT PRELIMINARY; PRT; 236 AA.
AC Q6P771;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE SERTA domain containing 1.
GN Name=SerTad1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.

```

RC TISSUE=prostate; PMID=12477932; DOI=10.1073/pnas.242603899;
RX STRAUSBERG R.L., FELINGOLD E.A., GROUSE L.H., DERGE J.G.,
RA KLAUSNER R.D., COLLINS F.S., WAGNER L., SHENMEN C.M., SCHULER G.D.,
RA ALTSCHUL S.F., ZEEBERG B., BUETOW K.H., SCHAEFER C.F., BHAT N.K.,
RA HOPKINS R.F., JORDAN H., MOORE T., MAX S.I., WANG J., HSEH F.,
RA DIACHENKO L., MARUSINA K., FARMER A.A., RUBIN G.M., HONG L.,
RA STAPLETON M., SOARES M.B., DONALDO M.F., CASAVANT T.L., SCHEETZ T.E.,
RA BROWNSTEIN M.J., USADIN T.B., TOSHIYUKI S., CARNINCI P., PRANGE C.,
RA RAHA S.S., LOQUELLANO N.A., PETERS G.J., ABRAMSON R.D., MULLAHY S.J.,
RA BOSAK S.A., MCEWAN P.U., MCKERNAN K.J., MALEK J.A., GUMARATNE P.H.,
RA RICHARDS S., WORLEY K.C., HALE S., GARCIA A.M., GAY L.J., HULYK S.W.,
RA VILLALON D.K., MUZNY D.M., SODERGREN E.J., LU X., GIBBS R.A.,
RA FAHEY J., HELTON E., KETTEMAN M., MADAN A., RODRIGUES S., SANCHEZ A.,
RA WHITING M., MADAN A., YOUNG A.C., SHEVCHENKO Y., BOUFFARD G.G.,
RA BLAKESLEY R.W., TOUCHMAN J.W., GREEN E.D., DICKSON M.C.,
RA RODRIGUEZ A.C., GRIMWOOD J., SCHMUTZ J., MYERS R.M.,
RA BUTTERFIELD Y.S.N., KRZWIWINSKI M.I., SKALEKA U., SMAIUS D.E.,
RA SCHMERCH A., SCHEIN J.E., JONES S.J.M., MARRA M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Prostate;
RA Director MGC Project;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC061808; AAH61808.1; -; mRNA.
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA; 1.
SQ SEQUENCE 236 AA; 25419 MW; E8B7502E0473ED8B CRC64;

Query Match	20.1%;	Score 208;	DB 2;	Length 236;
Best Local Similarity	31.9%;	Pred. No. 5.4e-08;		
Matches 76;	Conservative 33;	Mismatches 67;	Indels 62;	Gaps 13;

```

QY      4 GLKRRKSDLEEEERWE-----W-----SPAGLSYQOALLRISLDKVQSRSLG 46
      ||||| ||||| | :||:| :||:| :||:|
Db      5 GLKRR-----REEETMESLSVDSWMLDQSRPAVAQTPATVAS--SSLFDLVSVKLHSLR 58
      ||||| ||||| | :||:| :||:| :||:|
QY      47 PRAPSLRRHYLIHNTLQQLQALRLAPAPALPEPL-----FLGEDFSLSATIG 96
      ||||| ||||| :||:| :||:| :||:|
Db      59 QSEPDRLHLVIVNTILRKIQASME--PTTVLPPEPLOPTAPSVADNFLSSSDAGLSASMA 116
      ||||| ||||| :||:| :||:| :||:|

```

```

QY      97 SLRLRLDTSMDGTEPPQNPTPLGLQNEVPP-----QPDVFLEAL--SSRYLDSGL 147
      ||:|:|      ||:|:|      ||:|:|      ||:|:|
Db     117 SLLEDSHIEDLNQVPOQA-----DEGPPGRSVGYCLPNLGLDLGPAATGCLLDGL 170

```

```

Qy      148 DDFPLDIDTSAVEKE---PARAPBP-PHNLFCAQSGMWW-----NELDHIMEITLGS 196
          : | | | | : | | | | | | | | | | | | | | | | | | | | | | |
Db      171 EGLFEDIDTSMYDSELMLPASEGLKVPEN---GPAKETPELDEAEIDYLMYLVGT 225

```

RESULT 9
Q9DCZ2 MOUSE
ID Q9DCZ2 MOUSE PRELIMINARY;
PRT; 244 AA.

NC 2001-01-01 (TRIMBLrel. 17, Created)
DT 01-JUN-2001 (TRIMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRIMBLrel. 26, Last annotation update)
DT 01-MAR-2004 (TRIMBLrel. 26, Last annotation update)
DE Mus musculus adult male kidney cDNA, RIKEN full-length enriched
library, clone:0610008D10 product:TRIP-BR1, full insert sequence.

OS Mus musculus (Mouse).
OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;

ACN
RC
RX
NP
NUCLEOTIDE SEQUENCE.
STRAIN=C57BL/6J; TISSUE=Kidney;
MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;

RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
Meth. Enzymol. 303:19-44 (1999).
RL

RN [2]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Kidney;
MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RX

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R...

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glisi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaïdo I., Pesole G., Quackenbush
RA Schriml L.M., Staubli F., Suzuki R., Tomlita M., Wagner L., Washio

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilmings

RA Wymshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsuki S.,
RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).

RN [3]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C3/BL/6J; TISSUE=Kidney;
RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT *Analysis of the mouse transcriptome based on functional annotation

RT 60,770 full-length cDNAs.[†]
RL Nature 420:563-573 (2002).

RN [4]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

RA Carminci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Ito
RA Komno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT	"Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes"
RT	

RL Genome Res. 10:1617-1630 (2000).
RN [5]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Kidney;

RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P

RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamuro T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK023224; BAB22013.1; -; mRNA.
DR Ensembl; ENSMUSG00000008384; Mus musculus.
DR MGI; MGI:1913438; Sertad1.
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA; 1.
SQ SEQUENCE 244 AA; 26068 MW; 0F464F6419D362A6 CRC64;

Query Match 16.1%; Score 166; DB 2; Length 244;
Best Local Similarity 25.9%; Pred. No. 9.2e-05;
Matches 67; Conservative 33; Mismatches 55; Indels 104; Gaps 13;
QY 4 GLKRKHSDLBEBBERWE-----W----SPAGLQS----YQALLRISLDKVQSLGPR 48
Db 5 GLKRN----GBEBETMEALSVDSCWLDPSHPAVAQTPPTVASSSLFDLSVVKLHSLROS 60
QY 49 APSLRHVLINHTLQQLQALRLAPALPPEPL-----FLGEBDFSLSATIGSI 98
Db 61 EPDRLHLVLVNTLRRIQASME--PAPVLPEPIQPPAPSVADSLASSDAGLSASMASL 118
QY 99 LRELDTSMGTPEPQ-----NPVTP-----LG-----LONEV 125
Db 119 LEDLNHTIEDLNQAPQPADGPPGRSIGISPNLGALDLGQPLAVCWTMDWRACLRYSI 178
QY 126 PP-----QPDPFVFLBALSSRYLGDSDLDFLIDPITSAVEKEPARAPPEPPHNLFCA 177
Db 179 PPCTTVNYGYQ-----LTVSSAAPENGPAK--EERP----- 208
QY 178 PGSEWNELDHIMEIILGS 196
Db 209 --ELDEAEFLTDLMDVLVGT 225

RESULT 10
Q4T2W4_TETNG
ID Q4T2W4_TETNG PRELIMINARY; PRT; 244 AA.
AC Q4T2W4;
DT 13-SEP-2005 (Tremblrel. 31, Created)
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Chromosome undetermined SCAF10161, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00008134001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Catcolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volf J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01010161; CAF92768.1; -; Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 244 AA; 26476 MW; 796E0C82A3644A05 CRC64;

Query Match 15.8%; Score 163.5; DB 2; Length 244;
Best Local Similarity 29.8%; Pred. No. 0.00014;
Matches 68; Conservative 27; Mismatches 86; Indels 47; Gaps 8;
QY 4 GLKRKHSDLBEB-----EERWESPAGL-----QSYQALLRISLDKVQSLGPRAPSL 52
Db 29 GYGRKWSCLEDEALPAAGDKERSE--PDGFLTGPPSSVSWLGLCLEKLYHYQTGVGLSL 86
QY 53 RRHVLINHTLQQLQALRLAPALPPEPLFLGEBDFSLSATIGSILRELDTSMGTPEP 112
Db 87 RRSVLLINHTLRQIQEIQSDGMGTCAPEILGAGVQTDSC-----PLRQDLPTVTCPRGRPA 141
QY 113 QNPVTPLGLQNEVPPQDPVFLBALSS-RYLGDSGLDDFFLIDITSAVEKE----- 162
Db 142 RR-----SSRKTRPQAATLFGDANAAGVLSLALDIDFEDIDTSMYESSDLPANVAG 196
QY 163 ---PARAPPEPPHNLFCAGS-----WENELDHIMEIILGS 196
Db 197 SLMPVSISLWADEDLKVCSGSHASAGSLQSCRMDLNELDQIMEILVNS 244

RESULT 11
Q4S6F1_TETNG
ID Q4S6F1_TETNG PRELIMINARY; PRT; 275 AA.
AC Q4S6F1;
DT 13-SEP-2005 (Tremblrel. 31, Created)
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Chromosome 10 SCAF14728, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00023321001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Catcolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,

RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volf J.N., Guigo R., Zody R., Mestrov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype";
RL Nature 431:946-957(2004).

RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01014728; CAG03781.1; -; Genomic_DNA.
FT NON TER 275
SQ SEQUENCE 275 AA; 29572 MW; 9CF6A854C13C1349 CRC64;

Query Match 14.0%; Score 145; DB 2; Length 275;
Best Local Similarity 24.9%; Pred.No.0.0043;
Matches 72; Conservative 35; Mismatches 66; Indels 116; Gaps 12;

QY 4 GUKRKHSDLEEBEERWESPAGLQ-----SY---QQAALRISLDKVRSL 45
Db 5 GTRKRKFSADAGEBAA---APSGGQPPAATSAARTLSSSYSLQROSLDMSLRKQLCH 60
QY 46 GPRAPSLRHHVLIHNTLQQLQ-----AALRLAPAPALPPEPLFGSEDFSL 92
Db 61 MUYEPNLCRSVLIAANTVRQIQEEMTQDGTWQIMTKALAAQCPA-----DRLVA 109
QY 93 ATIGSILRELDTSMDGTEPPQNPVTPGLQ-----NEVPPQPD-----PYFLFA 136
Db 110 SEV--LCRQTDAAPAGQSPK-PFSVGLSEGYHSEEVWEGDMEPEVTMTSLSPVSPQL 166
QY 137 LSGRYLG-----DSGLDDFFLDIDTSAVERKEP 163
Db 167 SSASYLGPGFGMPCWEEBEDVTGEQVFTFEIKHPAPPPDPALELFSDVDPSTYYDLDT 226
QY 164 A-----RAPPEPHNLFCAPGSW-----EWNELDHIMEIIL 194
Db 227 VLTGMQSAPIKMGPYDLLESLSHGPTALSSSSQCRSLNELDHIMEIIV 275

RESULT 12
QSRK3_BRARE
ID QSRK3_BRARE PRELIMINARY; *PRT; 229 AA.
AC QSRK3;
DT 01-FEB-2005 (Tremblrel. 29, Created)
DT 01-FEB-2005 (Tremblrel. 29, Last sequence update)
DT 01-FEB-2005 (Tremblrel. 29, Last annotation update)
DE Novel protein (Fragment).
GN Name=si:key-177p2.6; Synonyms=OTDARP0000005081;
ORFNames=DKEY-177P2.6-001;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lovell J.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX248515; CAI20741.1; -; Genomic_DNA.
DR ZFIN; ZDB-GENE-030131-5926; si:key-177p2.6.
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA; 1.
FT NON TER 229
SQ SEQUENCE 229 AA; 25097 MW; F7936F752E74AA1F CRC64;

Query Match 14.0%; Score 144.5; DB 2; Length 229;
Best Local Similarity 29.2%; Pred.No.0.0038;
Matches 57; Conservative 14; Mismatches 61; Indels 63; Gaps 6;

QY 4 GUKRKHSDLEEBEERWESPAGLQSY---QQAALRISLDKVRSLGPRAPSLRHHVLIHN 60
Db 57 GVKRKLSTCED-----PAQDLPYPQQRQLVLDICDKLQSCQRRAPFSLRSHVLIAN 108
QY 61 TLQQLQAALR-----LAPA-----PALPPEPL----- 82
Db 109 TLRQIQEEMRQEGETCLPPAVLGPSPHIIQTPRHVPELPVPLDCPAPLTGALSPSPFLM 168
QY 83 FLGEEDFSLATIGSILRELDTSMDGTEPPQNPVTPGLQNEVPPQPDVFLALSSRYL 142
Db 169 TADEEELQGTENETLLPSLAGEDNTKSSDLLFGSFEITN-----STSYL 214
QY 143 GDSGLDDFFLDIDTS 157
Db 215 TDLALDDIFEDIDTS 229

RESULT 13
SRTD2_MOUSE
ID SRTD2_MOUSE STANDARD; PRT; 309 AA.
AC Q9JUG5; Q8C609; Q91WL3; Q925E5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE SERTA domain-containing protein 2 (Transcriptional regulator
interacting with the PHD-bromodomain 2) (TRIP-Br2).
GN Name=Sertad2; Synonyms=Klaa0127; ORFNames=MNCB-1504;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,
RA Hashimoto K.;
RT "Isolation of full-length cDNA clones from mouse brain cDNA library
made by oligo-capping method."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.B., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Peccole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Walestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shireki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 129-309.
RX MEDLINE=21331173; PubMed=11331592; DOI=10.1093/emboj/20.9.2273;
RA Hsu S.-I., Yang C.M., Sim K.G., Hentschel D.M., O'Leary E.,
RA Bonventre J.V.;
RT "TRIP-Br: a novel family of PHD zinc finger- and bromodomain-
RT interacting proteins that regulate the transcriptional activity of
RT E2F-1/DP-1.";
RL EMBO J. 20:2273-2285 (2001).
CC -1- FUNCTION: Acts at E2F-responsive promoters to integrate signals
CC provided by PHD- and/or bromodomain-containing transcription
CC factors (By similarity).
CC -1- SIMILARITY: Contains 1 SERTA domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AB041541; BAA95026.1; -; mRNA.
DR EMBL; AK076787; BAC36480.1; -; mRNA.
DR EMBL; BC014726; AAH14726.1; -; mRNA.
DR EMBL; AF366403; AAK52832.1; -; mRNA.
DR Ensembl; ENSMUSG0000049800; Mus musculus.
DR MGI; MGI:1931026; Sertad2.
DR InterPro; IPR009263; SERTA.
DR Pfam; PF06031; SERTA; 1.
DR PROSITE; PSS1053; SERTA; 1.
KW Transcription; Transcription regulation.
FT DOMAIN 33 80 SERTA.
FT CONFLICT 138 140 CTL -> FTP (in Ref. 4).
FT CONFLICT 142 142 A -> S (in Ref. 3).
FT CONFLICT 145 145 P -> S (in Ref. 4).
FT CONFLICT 234 234 T -> A (in Ref. 1).
SQ SEQUENCE 309 AA; 33312 MW; D4178688F0DF8F00 CRC64;
Query Match 13.6%; Score 140.5; DB 1; Length 309;
Best Local Similarity 24.8%; Pred. No. 0.011;
Matches 77; Conservative 29; Mismatches 85; Indels 119; Gaps 13;

QY 87 -----EDFSLATIGSILRELDTSMDG 108
Db 121 ACITPASLLEDNDTFTCTLQAVHPAPTRLSSAALPAEKDSFSSALDEI-EELCPSTST 179
QY 109 TEPPQN--PVTPLGLQNEVPPQ-----PDPVFLEAL-----SSRYLGDSGLD 148
Db 180 TEAAHTAAPGPKGTSSSESVQKPEGBEGRTDSDRFMDSLPGNFETTSTGFLTDLTLD 239
QY 149 D-PFLDIDTSAVEKEPA-----RAPPE-----PPHNLFCAPG---SWEMNEL 186
Db 240 DILFADIDTSMYDFDPCTASGTASKMAPVSADDLTKTLAPYSNQPVAPSPQFKMDLTTEL 299
QY 187 DHIMEIILGS 196
Db 300 DHIMEVLVGS 309
RESULT 14
Q5SSE5_MOUSE PRELIMINARY; PRT; 309 AA.
ID Q5SSE5_MOUSE PRELIMINARY;
AC Q5SSE5;
DT 10-MAY-2005 (Tremblrel. 30, Created)
DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)
DE Novel protein similar to cell division cycle associated 4 Cdca4.
GN Name=RP23-452C23.1; ORFNames=RP23-452C23.1-001;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Beasley H.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL663115; CAI25740.1; -; Genomic_DNA.
KW Cell division.
SQ SEQUENCE 309 AA; 33312 MW; D4178688F0DF8F00 CRC64;
Query Match 13.6%; Score 140.5; DB 2; Length 309;
Best Local Similarity 24.8%; Pred. No. 0.011;
Matches 77; Conservative 29; Mismatches 85; Indels 119; Gaps 13;
QY 2 EGGLRKHSDLBEEBERWESPAGLQS-----YQQALLRISLDKVQRLSGPRAPSLRRH 55
Db 4 KGG-KRKPEDEHEDGLEGKIVSPSDGPRSVSYTLQRTIFNISLMKLYNHRPLTEPSLQKT 62
QY 56 VLIHNTLQQLQALRL-----APAPALPPEPLFLGE----- 86
Db 63 VLINMMLRRIQELKQEGSLRAPFTPSSQPSNSLSDSYQEAAPPA--PHPCDLGSTTPLE 120
QY 87 -----EDFSLATIGSILRELDTSMDG 108
Db 121 ACITPASLLEDNDTFTCTLQAVHPAPTRLSSAALPAEKDSFSSALDEI-EELCPSTST 179
QY 109 TEPPQN--PVTPLGLQNEVPPQ-----PDPVFLEAL-----SSRYLGDSGLD 148
Db 180 TEAAHTAAPGPKGTSSSESVQKPEGBEGRTDSDRFMDSLPGNFETTSTGFLTDLTLD 239
QY 149 D-PFLDIDTSAVEKEPA-----RAPPE-----PPHNLFCAPG---SWEMNEL 186
Db 240 DILFADIDTSMYDFDPCTASGTASKMAPVSADDLTKTLAPYSNQPVAPSPQFKMDLTTEL 299
QY 187 DHIMEIILGS 196
Db 300 DHIMEVLVGS 309
RESULT 15
Q6GM81_XENLA
ID Q6GM81_XENLA PRELIMINARY; PRT; 248 AA.
AC Q6GM81;

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 10:01:21 ; Search time 49 Seconds
(without alignments)
330.703 Million cell updates/sec

Title: US-10-069-386A-2
Perfect score: 1033
Sequence: 1 MEGGLKRKHSDLKEEERWE.....APGSWEWNLDHIMEIILGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	136	13.2	237	2	US-10-076-069-2 Sequence 2, Appli
2	105	10.2	241	2	US-10-076-069-4 Sequence 4, Appli
3	98	9.5	562	2	US-10-104-047-3214 Sequence 3214, Ap
4	95.5	9.2	529	2	US-09-949-016-8918 Sequence 8918, Ap
5	94	9.1	433	2	US-08-259-451-3 Sequence 3, Appli
6	93.5	9.1	1734	2	US-09-862-027-81 Sequence 81, Appl
7	93	9.0	433	1	US-07-672-483-4 Sequence 4, Appli
8	93	9.0	2618	2	US-09-413-814-28 Sequence 28, Appl
9	92.5	9.0	113	2	US-10-076-069-6 Sequence 6, Appli
10	92.5	9.0	432	4	PCT-US95-04910-13 Sequence 13, Appl
11	92	8.9	379	2	US-09-949-016-10257 Sequence 10257, A
12	90.5	8.8	370	2	US-09-377-285B-16 Sequence 16, Appl
13	90.5	8.8	370	2	US-10-192-381-16 Sequence 16, Appl
14	90.5	8.8	824	2	US-09-949-002-312 Sequence 312, App
15	90.5	8.8	947	1	US-08-887-518-2 Sequence 2, Appli
16	90.5	8.8	947	1	US-09-023-321-2 Sequence 2, Appli
17	90.5	8.8	947	1	US-09-032-475-2 Sequence 2, Appli
18	90.5	8.8	947	2	US-09-257-703-1 Sequence 1, Appli
19	90.5	8.8	947	2	US-09-871-889A-1 Sequence 1, Appli
20	90.5	8.8	947	2	US-09-981-397A-18 Sequence 18, Appl
21	90.5	8.8	1027	2	US-09-502-540-11750 Sequence 11750, A
22	89.5	8.7	215	2	US-08-778-717-9 Sequence 9, Appli
23	89.5	8.7	611	2	US-09-949-016-7139 Sequence 7139, Ap
24	88.5	8.6	897	2	US-09-849-602-18 Sequence 18, Appl
25	88	8.5	336	2	US-10-104-047-2029 Sequence 2029, Ap
26	88	8.5	1130	2	US-09-976-594-280 Sequence 280, App
27	87.5	8.5	709	2	US-10-118-328-4 Sequence 4, Appli

28	86.5	8.4	136	2	US-08-259-451-5	Sequence 5, Appli
29	86	8.3	337	2	US-10-144-929-97	Sequence 97, Appl
30	85	8.2	1148	2	US-09-949-016-6798	Sequence 6798, Ap
31	84	8.1	337	2	US-09-543-681A-7444	Sequence 7444, Ap
32	84	8.1	514	2	US-09-252-991A-25281	Sequence 25281, A
33	84	8.1	628	2	US-09-345-473E-48	Sequence 48, Appl
34	84	8.1	628	2	US-09-862-027-48	Sequence 48, Appl
35	84	8.1	675	2	US-09-949-016-11490	Sequence 11490, A
36	84	8.1	1015	2	US-09-949-016-6276	Sequence 6276, Ap
37	84	8.1	2152	2	US-09-036-987A-3	Sequence 3, Appli
38	84	8.1	2152	2	US-09-370-700-3	Sequence 3, Appli
39	84	8.1	2152	2	US-09-603-207-3	Sequence 3, Appli
40	83.5	8.1	1728	2	US-09-949-002-532	Sequence 532, App
41	83	8.0	969	1	US-08-548-159-1	Sequence 1, Appli
42	83	8.0	986	1	US-08-548-159-3	Sequence 3, Appli
43	83	8.0	1012	2	US-08-811-481-16	Sequence 16, Appl
44	83	8.0	1012	2	US-09-876-527-16	Sequence 16, Appl
45	82.5	8.0	1130	2	US-09-538-092-834	Sequence 834, App

ALIGNMENTS

RESULT 1
US-10-076-069-2
; Sequence 2, Application US/10076069
; Patent No. 6872812
; GENERAL INFORMATION:
; APPLICANT: JURECIC, ROLAND
; APPLICANT: NACHTMAN, RONALD
; TITLE OF INVENTION: HEPP, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELOPMENT
; FILE REFERENCE: 39532-176599
; CURRENT APPLICATION NUMBER: US/10/076, 069
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/268, 923
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-076-069-2

Query Match	Best Local Similarity	Score	DB 2;	Length	Matches	Conservative	Pred. No.	Mismatches	Indels	Gaps
QY	4	GLKRGKHSDLKEEERWWSPA-GLQSYQALLRISLDKVGQSLGPRAPSLRRHVLINHTL	62	63	QQLQALR-----LAP--APALPEPFLFGEEDFSLSATIGSLRELDTSMDGTEPPQ	113	13.2%;	26.1%;	86;	60;
DB	5	GLKRGKGDQEEGVEGFTVPYSYLQ--RQSLDMSLVKLQCLHMLVBPNLCSRVLIAHTV	62	63	RQIQEEMSQDGVWHGMAPQNVDRAPVERLVSTE--ILCRTVRGAEEHHPAPELEDAPLQ	119	32;	3.2e-05;	86;	9;
QY	114	NPVTPLGLQNEVPQPDP-----VFLEALSSRYLGDSGLDDFF	151	120	NSVSELPIVGSAPGRNPQSSLMEMDSPQENRGSPQKSLDQIFETLENK--NSSSVEBLF	177	136;	26.1%;	86;	60;
DB	152	LDIDTSAVEKEPAR-----APPEPPHNLFCAPGSWEWNLDHIMEII	193	178	SDVDSYYDDLTDTLTGMMSGTKSLCNGLEGFAAATPPPSGTCKS---DLAELDHVEIL	234	136;	26.1%;	86;	60;
QY	194	L 194		235	V 235		136;	26.1%;	86;	60;
DB							136;	26.1%;	86;	60;

RESULT 2
US-10-076-069-4
; Sequence 4, Application US/10076069
; Patent No. 6872812


```

; TITLE OF INVENTION: HTLV-I/II/III Compositions
; TITLE OF INVENTION: and Assays for Detecting HTLV Infection
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: One Abbott Park Road
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60064
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; MEDIUM TYPE: storage
; COMPUTER: IBM
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/259,451
; FILING DATE: 20-JUN-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/086,415
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Daniel W. Collins
; REGISTRATION NUMBER: 31,912
; REFERENCE/DOCKET NUMBER: 5381.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708) 937-6365
; TELEFAX: (708) 938-2623
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
;
US-08-259-451-3

```

```

Query Match          9.1%; Score 94; DB 2; Length 433;
Best Local Similarity 23.1%; Pred. No. 0.78;
Matches 40; Conservative 24; Mismatches 67; Indels 42; Gaps 7;

QY 47 PRAPSLRRHVLIHNTLQOLQALRLAPAPA-----LPPBPLFGBEDFSLSA 93
   ||| | : | : ||| ||| | : | :
DB 13 PKAP---RGLSTHWNLFQAAYRLQPGPSDFDQQLRFLKLAKTPIWLNPIIDYSLA 69
   ||| | : | : ||| ||| | : | :
QY 94 TI-----GSILRELDTSMDGTEPPQNPVTPLG--LQNEVPQPPDPVFLBA-LSSRYLG 143
   : : : : : | : | : | : | : | : | :
DB 70 SLIPKGYPRVVEIINILVKNQVSPSAPAPVPTPICPTTTPPPPPPSPEAHVPPPY-- 127
   : : : : : | : | : | : | : | : | :
QY 144 DSGLDFFLDIDTSAVEKEPARAPPEPPHNLFCAPGSWEWNELDHIMEIILGS 196
   : : : : : | : | : | : | : | : | :
DB 128 -----VEPTTQCFPIILHPPGAP-----SAHRPMQMKDLQAIKQBVSSS 165
   : : : : : | : | : | : | : | : | :

RESULT 6
US-09-862-027-81
; Sequence 81, Application US/09862027
; Patent No. 6858418
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. 6858418el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862, 027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345, 473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 1734
; TYPE: PRT
; ORGANISM: Mus musculus

```

```
US-09-862-027-81
```

	Query Match Best Local Similarity Matches	9.1%; 26.3%; 51;	Score 93.5; Pred. No. 5.3; Conservative	DB 2; Mismatched 23;	Length 1734; 71;	Indels 49; Gaps	8;
QY	25 GLQSYQQAL-----LRISLDKVVQRSLGPRAPIRRHYLIHNTLLOQLQAALRLAPARALPP	79					
Dd	1308 GSQSFPFKHLSPPLGRQLSRPKSABPFRSPLKR-----VQSAEKLAALLAAAE	1357					
QY	80 EPLFLGEEDFSLSATIGSILRELDTSMGTBPONPVTPGLQNVP--PODPVFLEA	136					
Dd	1358 KKL-APS RKHSIDLPHGELKKEL-----TPREASPLEVGTRSVLSGKGPLPGKVLP	1410					
QY	137 LSSRYLG-----DSGLDFFLLIDITSAVEKEPARAP--DEPPHN	173					
Dd	1411 APSRALGTLRQDRARRRESLQKAIRREVDSSEDDTDEEPENSQATQEPRLSPEASHN	1470					
QY	174 LFCAPGSWEWNELD	187					
Dd	1471 LL-PKSGEGTEED	1483					

```

RESULT 7
US-07-672-483-4
; Sequence 4, Application US/07672483
; Patent No. 5359029
; GENERAL INFORMATION:
; APPLICANT: LACROIX, Martial
; APPLICANT: ZREIN, Maan
; TITLE OF INVENTION: PEPTIDES AND ANALOGUES AND MIXTURES
; TITLE OF INVENTION: THEREOF FOR DETECTING ANTIBODIES TO HTLV-I AND HTLV-II
; TITLE OF INVENTION: VIRUSES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/672,483
; FILING DATE: 19910302
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/554,258
; FILING DATE: 18-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY Jr, James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: IAF8 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0674
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-07-672-483-4

```

```

Query Match          9.0%; Score 93; DB 1; Length 433;
Best Local Similarity 23.1%; Pred. No. 0.97;
Matches 40; Conservative 24; Mismatches 67; Indels 42; Gaps 7;

```

```
OY   PRAPSLRRHVLINHTLQOLQAALRLAPAPA-----LPPEPLFLGEBDFLSLA 93
      ||| : | : ||||| : : ||| : |||
Db   PKAP---RGLSTHHMLNLFQAAYRLQPBPSPDFEQQLRRFLKALKTPIMWNPIDYSISLA 69
      ||| : | : ||||| : : ||| : |||
OY   TI-----GSILRELDTSMDGTBPQNPTPLG---LQNEVPPQDPDVLEA-ISSRYLG 143
      :: : : : : : : ||| : : : :
Db   SLIPKGYPRVVEIINILVKNOVSPSAPAAVPPTPICPTTTPRPSPPEAHVPPPY-- 127
      : : : : : : : : : : : :
OY   DSGLDFFLLDIDTSAVEKEPARAPPEPPHNLFCAGSWENELDHIMEILGS 196
      :: : : | : | : : : : : : :
Db   -----VEPTTQCFFPIHLHPGAP-----SAHRPWOMKDLOAIKQEVSSS 166
      : : : : : : : : : : : :
```

RESULT 8
US-09-413-814-28
; Sequence 28, Application US/09413814

```

; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloecker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; TITLE OF INVENTION: heteropolyketide compounds
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413, 814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 2618
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-28

```

	Query Match	9.0%	Score 93;	DB 2;	length 2618;
	Best Local Similarity	25.5%;	Pred No. 10;		
	Matches	55; Conservative	18; Mismatches	79; Indels	64; Gaps 11;
Qy	31 QALLR---ISLD-----KVRSLGPRAP--SLRRHVLINHTLOOLQA	67			
Dd	974 QALLRREALIGLDEPFQAGNSFGILRLAKLESFAFGKSPFITDLFQHTSIRSQAEMLSG	1033			
Qy	68 ALRLAP-APALPPEPLFIQEEDFSUSATIGSILRELDTSMDGTERRPQ---NPVTPIGLQN	123			
Dd	1034 SSVEAPLAGAVPPQPRAAAQVASSAAXSPGERGAATSSGLTIAQPPQHFRPIAVIGLAG	1093			
Qy	124 EVFPQPD-PVFLLBAL-----SSRYLGDSGLD-----DFFLD	153			
Dd	1094 RPPAAPDLDAFLELLTEGRCCGIRFFSOAELRDQGLDANRIACHNVYPAKGFLDRADHF-D	1152			
Qy	154 IDTSAVEKEPAR-APPEPPHNLFCAFGSWENELDH	188			
Dd	1153 ADFFGIIPRD AEITDPQIRLLLEC-----WNAL E H	1183			

RESULT 9
US-10-076-069-6
; Sequence 6, Application US/10076069
; Patent No. 6872812
; GENERAL INFORMATION:
; APPLICANT: JURECIC, ROLAND
; APPLICANT: NACHTMAN, RONALD
; TITLE OF INVENTION: HEPV, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELOPMENT

```

; FILE REFERENCE: 39532-176599
; CURRENT APPLICATION NUMBER: US/10/076,069
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/268,923
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-076-069-6

```

Query Match	9.04%	Score 92.5;	DB 2;	Length 113;
Best Local Similarity	45.34%	Pred. No. 0.19;		
Matches	29;	Conservative	11;	Mismatches 21;
				Indels 3;
				Gaps 2;

[illegible]

RESULT 10
PCT-US95-04910-13
: Sequence 13, Application PC/TUS9504910

GENERAL INFORMATION:
 APPLICANT: The Government of the United States of America as represented
 APPLICANT: by the Secretary, Department of Health and Human Services
 TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A NOVEL PRIMATE T-CELL
 TITLE OF INVENTION: LYMPHOTROPIC VIRUS AND THE USE OF THIS VIRUS
 TITLE OF INVENTION: OR COMPONENTS THEREOF IN DIAGNOSTIC ASSAYS
 TITLE OF INVENTION: AND VACCINES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & PINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1

APPLICATION NUMBER: PCT/US95/04910
FILING DATE: 21-APR-1995
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US08/231,526
FILING DATE: 22-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4125PCT
TELECOMMUNICATION INFORMATION:

TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acids
STRANDEDNESS: unknown
TOPOLOGY: unknown


```

; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/097,334
; PRIOR FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-192-381-16

```

Query Match	8.8%	Score 90.5;	DB 2;	Length 370;
Best Local Similarity	23.5%;	Pred. No. 1.4;		
Matches 43;	Conservative 29;	Mismatches 62;	Indels 49;	Gaps 8;

```

QY      33 LIRISLDKV-QRSLGPRAPSLRRHVLIHNTLQQLQAALRLAPAPALPP-----EPLFLG 85
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      102 MNMNHLEKVARREIGTLA-----TVRRLPSQKVIPEESLPLTPYCRKPL--- 147

QY      86 EEDFSLSATIGSILRELDTSMDGT-----EPPQNVPVTPLGLQNEVPPQ-PDPVFLAL-- 137
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      148 --NFACLDDVGHGKDLSTQLSRTGTLRSKSIKAPATPPASATLGRPPRIPEPQQLPAVPD 205

QY      138 -----SSRYLGDSGLDDFLDIDITSANVEKEPARAPPEP-----PHNLFCAFG 179
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      206 GKLSAASSVSVSLASAGSAGASGIPQSKGVAPATPPPPPIAPVTPPPPPPLPAIIFLLPP 265

QY      180 SWE 182
      |
Db      266 PME 268

```

```

RESULT 14
US-09-949-002-312
; Sequence 312, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 312
; LENGTH: 824
; TYPE: PRT
; ORGANISM: Human
US-09-949-002-312

```

Query Match	8.8%	Score 90.5	DB 2	Length 824
Best Local Similarity	25.0%	Pred. No. 3.9		
Matches 39	Conservative 24	Mismatches 42	Indels 51	Gaps 8
QY	24	AGIOSYQOALLRISLDKVRSLGPRAPSRRHVLIHNTLQOLQALRL-----APAPALP	78	
Db	673	AGIIVYRKARSRI-----LSRNVAPKTTGSRNPLFH-----QAASRVPAKGAPAPSRG	722	
QY	79	PEPFLGEEEDFSLSATIGSILRELDTSMDGTEPPQNVPVTLGLQNEVPPQDEVFLEALS	138	
Db	723	PQEL-----VPTTHPG-QPARHPASSVALKR--PPRAPVTVSS--	758	
QY	139	SRYLDSGLDDFFLDIDITSAVEKE---PARAPPEPP	171	
Db	759	-----PPFPVPVYTRQAPKQVIKPTFADPPVP	785	

RESULT 15
US-08-887-518-2

```

; Sequence 2, Application US/08887518
; Patent No. 5843721
;
; GENERAL INFORMATION:
; APPLICANT: Rothe, Mike
; APPLICANT: Wu, Lin
; TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,518
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 947 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-08-887-518-2

```

Query Match	8.8%;	Score 90.5;	DB 1;	Length 947;
Best Local Similarity	25.8%;	Pred. No. 4.7;		
Matches	54;	Conservative 17;	Mismatches 87;	Indels 51; Gaps 10;
QY	3	GGLKRKHSDL EE E E E E W E W S P A G L Q S Y Q A A L R I S L D K V Q R S L G P R A P S L R R H V L I H N T L	62	
Db	659	GGLK--SPWRGEYKEPRHPPPNQANYHQ-----TLHAQPRELSPRAGPRPAEETGRA	710	
QY	63	QQLQAALRIAPAPALPPE-----PLFLGEEDFSLATIGSILRELDTSMGTEPPONPYT	117	
Db	711	PKLQ-----PPLPPEPEPNKSPPLTISKEE-----SGMWEPLPLSSLRPAPARNPSS	758	
QY	118	PLGLQNEVPQ-----PDVFLEALSSRY-----LGSGLDDFFLDIDITSAVEKEPA	164	
Db	759	P-EKATVPEQELQQLTELEFLNLSLQPFSLAEQELISCLSIDSLSDDS--EKNPSS	814	
QY	165	RAPPEPPHNLFCAPGSW-----EWN	184	
Db	815	KASQSSRDTLSSGVHSSWSQAEARSSSWN	843	

Search completed: February 6, 2006, 10:02:50
Job time : 50 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 10:13:16 ; Search time 171 Seconds
(without alignments)
478.915 Million cell updates/sec

Title: US-10-069-386A-2
Perfect score: 1033
Sequence: 1 MEGGLKRKHSDLKEEERWE.....APGSWEWELDHIMEIILGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 segs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	791	76.6	157	4	US-10-106-698-5516 Sequence 5516, Ap
2	645.5	62.5	142	4	US-10-029-386-33709 Sequence 33709, A
3	402	38.9	111	4	US-10-029-386-33708 Sequence 33708, A
4	217	21.0	236	6	US-11-095-870-15 Sequence 15, Appl
5	216	20.9	236	3	US-09-978-360A-568 Sequence 568, App
6	213	20.6	236	3	US-09-731-872-412 Sequence 412, App
7	213	20.6	236	3	US-09-876-997-412 Sequence 412, App
8	213	20.6	236	5	US-10-643-836-412 Sequence 412, App
9	213	20.6	236	6	US-11-095-870-14 Sequence 14, Appl
10	213	20.6	278	4	US-10-296-115-1065 Sequence 1065, Ap
11	205	19.8	222	4	US-10-029-386-34001 Sequence 34001, A
12	136	13.2	237	4	US-10-076-069-2 Sequence 2, Appli
13	125	12.1	314	4	US-10-170-385-31 Sequence 31, Appl
14	125	12.1	314	4	US-10-755-889-627 Sequence 627, App
15	125	12.1	314	6	US-11-095-870-12 Sequence 12, Appl
16	106	10.3	254	3	US-09-925-302-596 Sequence 596, App
17	106	10.3	254	3	US-09-925-302-596 Sequence 596, App
18	106	10.3	254	4	US-10-106-698-4456 Sequence 4456, Ap
19	105	10.2	241	4	US-10-076-069-4 Sequence 4, Appli
20	105	10.2	241	5	US-10-965-898-51 Sequence 51, Appl
21	104.5	10.1	740	4	US-10-276-774-2083 Sequence 2083, Ap
22	104	10.1	322	5	US-10-450-763-53227 Sequence 53227, A
23	103.5	10.0	181	6	US-11-095-870-13 Sequence 13, Appl
24	98	9.5	562	4	US-10-104-047-3214 Sequence 3214, Ap
25	97	9.4	803	4	US-10-369-493-2536 Sequence 2536, Ap
26	97	9.4	803	5	US-10-732-923-7169 Sequence 7169, Ap
27	97	9.4	1099	4	US-10-425-115-195088 Sequence 195088,

28	96	9.3	297	4	US-10-425-114-64002	Sequence 64002, A
29	96	9.3	817	4	US-10-369-493-1813	Sequence 1813, Ap
30	95.5	9.2	1228	4	US-10-437-963-188300	Sequence 188300,
31	95	9.2	278	5	US-10-739-930-8565	Sequence 8565, Ap
32	95	9.2	713	4	US-10-437-963-137248	Sequence 137248,
33	94	9.1	433	2	US-08-259-451-3	Sequence 3, Appli
34	94	9.1	433	4	US-10-224-999A-3472	Sequence 3472, Ap
35	93.5	9.1	256	4	US-10-421-138A-312	Sequence 312, App
36	93.5	9.1	256	4	US-10-374-780A-1247	Sequence 1247, Ap
37	93.5	9.1	429	6	US-11-097-143-4629	Sequence 4629, Ap
38	93.5	9.1	663	4	US-10-755-889-480	Sequence 480, App
39	93.5	9.1	663	3	US-10-756-149-4873	Sequence 4873, Ap
40	93.5	9.1	1734	5	US-09-862-027-81	Sequence 81, Appl
41	93.5	9.1	1734	4	US-10-042-865-82	Sequence 82, Appl
42	93.5	9.1	1734	5	US-10-989-228-81	Sequence 81, Appl
43	93	9.0	157	4	US-10-767-701-54984	Sequence 54984, A
44	92.5	9.0	113	4	US-10-076-069-6	Sequence 6, Appli
45	92	8.9	1103	4	US-10-437-963-166325	Sequence 166325,

ALIGNMENTS

RESULT 1
US-10-106-698-5516
; Sequence 5516, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patentin Ver. 3.0
; SEQ ID NO 5516
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (132)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5516

Query Match 76.6%; Score 791; DB 4; Length 157;
Best Local Similarity 98.0%; Pred. No. 1.2e-63;
Matches 150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 44 SLGPRAPSLRRHVLHNTLQQLAALRLAPAPALPREPLFGEEDFSLSATIGSILRELD 103
:|||||
Db 5 ALGPRAPSLRRHVLHNTLQQLAALRLAPAPALPREPLFGEEDFSLSATIGSILRELD 64

QY 104 TSMGTEPPQNVTPLGLONEVPPQDPVFLEALSSRYLGDSGLDDFFLDIDTSAVERKEP 163
|||||
Db 65 TSMGTEPPQNVTPLGLONEVPPQDPVFLEALSSRYLGDSGLDDFFLDIDTSAVERKEP 124

QY 164 ARAPEPPPHNLFCAPGSWEWNELDHIMEIILGS 196
|||||
Db 125 ARAPEPPPHNLFCAPGSWEWNELDHIMEIILGS 157

RESULT 2
US-10-029-386-33709

```
; Sequence 33709, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33709
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010271.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
US-10-029-386-33709
```

Query Match 62.5%; Score 645.5; DB 4; Length 142;
Best Local Similarity 96.8%; Pred. No. 1.7e-50;
Matches 122; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```
OY 71 LAPAPLPEPLFLGBDFSLSATIGSILRELDTSMDGTEPPQNPTPLGLQNEVPPQPD 130
Db 18 LPPLPC-PPEPLFLGEEDFSLSATIGSILRELDTSMDGTEPPQNPTPLGLQNEVPPQPD 76
OY 131 PVFLALSSRYLGSDGLDDFLLDIDTSAVEKEPARAPPEPPHNLFCAPGSWENELDHIM 190
Db 77 PVFLALSSRYLGSDGLDDFLLDIDTSAVEKEPARAPPEPPHNLFCAPGSWENELDHIM 136
OY 191 EILGS 196
Db 137 EILGS 142
```

RESULT 3
US-10-029-386-33708
; Sequence 33708, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33708
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010271.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.64
US-10-029-386-33708

Query Match 38.9%; Score 402; DB 4; Length 111;
Best Local Similarity 89.9%; Pred. No. 1.5e-28;
Matches 80; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
OY 1 MEGGLKRKHSDLSEERWSPAGLSYQALRLISLDKQKSLGPRAPSLRRHVLIN 60
Db 3 MYGGLKRKHSDLSEERWSPAGLSYQALRLISLDKQKSLGPRAPSLRRHVLIN 62
OY 61 TLQQLQALRLAPAPALPPEPLFLGEEDF 89
```

Db 63 TLQQLQALRLAPAPALPPEPLPGRGCF 91

RESULT 4
US-11-095-870-15
; Sequence 15, Application US/11095870
; Publication No. US20050222034A1
; GENERAL INFORMATION:
; APPLICANT: Heu, Stephen I-Hong
; TITLE OF INVENTION: MODULATION OF TRIP-BR FUNCTION AND METHOD OF TREATING
; TITLE OF INVENTION: PROLIFERATIVE DISORDERS
; FILE REFERENCE: 93231-89
; CURRENT APPLICATION NUMBER: US/11/095,870
; CURRENT FILING DATE: 2005-03-31
; PRIOR APPLICATION NUMBER: US 60/557,697
; PRIOR FILING DATE: 2004-03-31
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: TRIP-Brl
US-11-095-870-15

Query Match 21.0%; Score 217; DB 6; Length 236;
Best Local Similarity 31.5%; Pred. No. 2.2e-11;
Matches 76; Conservative 33; Mismatches 64; Indels 68; Gaps 12;

```
OY 4 GLKRXHSDLEEEERWE-----W-----SPAGLOS-----YQALLRLISLDKQKSLGPR 48
Db 5 GLKRX-----RSEETMEALSVDSCWLDPSHPAVAGTPTVAASSLFDLSVVKLHSLROS 60
OY 49 APSLRHVLINHTLQQLQALRLAPAPALPPEPL-----FLGEEDFSLSATIGSI 98
Db 61 EPDLRLVLVNTLRIRIQASME--PAPVLPPEPIQPPAPPSVADSLASSDAGLSASMASL 118
OY 99 LRELDTSMDGTEPPQNPTPLGLQNEVPP-----QDPVFLAL--SSRYLGSDGLDD 149
Db 119 LEDLNHIBDLNQAQPQA-----DEGPRGSIGISPNLGAIDLGPATGCLDDGHEG 172
OY 150 FFLDIDTSAVEKE--PAR-----APPEPPHNLFCAPGSWENELDHIMEILG 195
Db 173 LFEDIDTSMYDSELWLPASEGLKPGPENGPAKEEPP-----ELDEAEILDYLMDEVLYG 224
OY 196 S 196
Db 225 T 225
```

RESULT 5
US-09-978-360A-568
; Sequence 568, Application US/09978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563


```
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNA5 ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/10/643,836
; PRIOR FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 412
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-643-836-412
```

```
Query Match      20.6%; Score 213; DB 5; Length 236;
Best Local Similarity 32.6%; Pred. No. 5.1e-11;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;
```

```
QY      4 GLKRKHSDLSEEE---ERWWSPPAGLSYQQA-----LLRISLDKVQRSIGPPAPS 51
          ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      5 GLKRKREEEKEPLAVDSW-WLDPGHATAVAQAAPPAVASSSLFDLSVLKLNHSLQOSEPD 63

QY      52 LRRHVLHNTLQOLAALRLAPALPPEPL-----FLGEBDFSLSATIGSLRE 101
          || : || : || : || : || : || : || : || : || : || : || :
Db      64 LRHLVLVNTLRRIQAS--MAPAALPVPSPPAAPSVADNLLASSDAALSASMASLLED 121

QY      102 LDTSMGTEPPQNPTPLGLQNEVPP-----QPDVFLAL--SSRYLGDSGLDDFFL 152
          | : : | : | : | : | : | : | : | : | : | : | : | : | :
Db      122 L-SHIEGLSQAPQ-----LADEGPPGRSIGGAAPSLGALDILGPATGCLDDGLEGLFE 175

QY      153 DIDTSAVEKE---PARAPPEP-PHNLFCAFGSWE-----WNELDHIMEITLGS 196
          ||||| : | : | : | : | : | : | : | : | : | : | : | :
Db      176 DIDTSMYDNELMAPASEGLKGPED--GPGKEAPELDEAEILDYLMVDLVGT 225
```

RESULT 9

```
US-11-095-870-14
; Sequence 14, Application US/11095870
; Publication No. US20050222034A1
; GENERAL INFORMATION:
; APPLICANT: Hsu, Stephen I-Hong
; TITLE OF INVENTION: MODULATION OF TRIP-BR FUNCTION AND METHOD OF TREATING
; TITLE OF INVENTION: PROLIFERATIVE DISORDERS
; FILE REFERENCE: 93231-89
; CURRENT APPLICATION NUMBER: US/11/095,870
; CURRENT FILING DATE: 2005-03-31
; PRIOR APPLICATION NUMBER: US 60/557,697
; PRIOR FILING DATE: 2004-03-31
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: TRIP-Br1
; US-11-095-870-14
```

```
Query Match      20.6%; Score 213; DB 6; Length 236;
Best Local Similarity 32.6%; Pred. No. 5.1e-11;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;
```

```
QY      4 GLKRKHSDLSEEE---ERWWSPPAGLSYQQA-----LLRISLDKVQRSIGPPAPS 51
          ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
```

```
Db      5 GLKRKREEEKEPLAVDSW-WLDPGHAAVAQAAPPAVASSSLFDLSVLKLNHSLQOSEPD 63

QY      52 LRRHVLHNTLQOLAALRLAPALPPEPL-----FLGEBDFSLSATIGSLRE 101
          || : || : || : || : || : || : || : || : || : || : || :
Db      64 LRHLVLVNTLRRIQAS--MAPAALPVPSPPAAPSVADNLLASSDAALSASMASLLED 121

QY      102 LDTSMGTEPPQNPTPLGLQNEVPP-----QPDVFLAL--SSRYLGDSGLDDFFL 152
          | : : | : | : | : | : | : | : | : | : | : | : | : | :
Db      122 L-SHIEGLSQAPQ-----LADEGPPGRSIGGAAPSLGALDILGPATGCLDDGLEGLFE 175

QY      153 DIDTSAVEKE---PARAPPEP-PHNLFCAFGSWE-----WNELDHIMEITLGS 196
          ||||| : | : | : | : | : | : | : | : | : | : | : | :
Db      176 DIDTSMYDNELMAPASEGLKGPED--GPGKEAPELDEAEILDYLMVDLVGT 225
```

RESULT 10

```
US-10-296-115-1065
; Sequence 1065, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hysq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1065
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-296-115-1065
```

```
Query Match      20.6%; Score 213; DB 4; Length 278;
Best Local Similarity 32.6%; Pred. No. 6.3e-11;
Matches 76; Conservative 36; Mismatches 69; Indels 52; Gaps 13;
```

```
QY      4 GLKRKHSDLSEEE---ERWWSPPAGLSYQQA-----LLRISLDKVQRSIGPPAPS 51
          ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      47 GLKRKREEEKEPLAVDSW-WLDPGHAAVAQAAPPAVASSSLFDLSVLKLNHSLQOSEPD 105

QY      52 LRRHVLHNTLQOLAALRLAPALPPEPL-----FLGEBDFSLSATIGSLRE 101
          || : || : || : || : || : || : || : || : || : || : || :
Db      106 LRHLVLVNTLRRIQAS--MAPAALPVPSPPAAPSVADNLLASSDAALSASMASLLED 163

QY      102 LDTSMGTEPPQNPTPLGLQNEVPP-----QPDVFLAL--SSRYLGDSGLDDFFL 152
          | : : | : | : | : | : | : | : | : | : | : | : | : | :
Db      164 L-SHIEGLSQAPQ-----LADEGPPGRSIGGAAPSLGALDILGPATGCLDDGLEGLFE 217

QY      153 DIDTSAVEKE---PARAPPEP-PHNLFCAFGSWE-----WNELDHIMEITLGS 196
          ||||| : | : | : | : | : | : | : | : | : | : | : | :
Db      218 DIDTSMYDNELMAPASEGLKGPED--GPGKEAPELDEAEILDYLMVDLVGT 267
```

RESULT 11

```
US-10-029-386-34001
; Sequence 34001, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34001
```

```

; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010271.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.46
; OTHER INFORMATION: SWISSPROT HIT: Q14140, EVALU8 2.70e-01
US-10-029-386-34001

```

Query Match	19.8%;	Score 205;	DB 4;	Length 222;
Best Local Similarity	32.6%;	Pred. No. 2.5e-10;		
Matches	75;	Conservative	34;	Mismatches 69;
			Indels	52;
			Gaps	13;

```
QY      4 GLKRKHSDLBEE-----ERWEMSPAGLOSYQA-----LIRISLDKVQRSLGPAPS   51
        ||||| : ||| : | | | : | : | : | |
Db      5 GLKRKRREEEKEPLAVDSW-WLDPGHTAVAQA PPAVASSSLFDLSVLKLHSHLQQSEPD   63
QY      52 LRHRVLIHNTLQOLAALRLAPAPALPEPL-----FLGBEDFSLSATIGSILRE   101
        ||| : ||| : ||| : | | | | | : | : | : | : | : | :
Db      64 LRLHLVVNTLRRIQAS--MAPAALBPVESPPAPASVADNLLASSDAALSAMASILED   121
QY      102 LDTSMDGTETPQNPTVTPLGLQNEVP-----QPDPVFLEAL--SSRYLGDSGLDDFFEL   152
         | : : | | | : | : | : | : | : | : | : | : | :
Db      122 L-SHIEGLSQARQP-----LADEGPPERSIGGAAPSLGALLDLPATGCCLDGEGLBE   175
QY      153 DIDTSAVEKE---PARAPPEP-PHNLFCAPGSWE-----WNELDHIMEIT   193
        ||||| : | | | : | | | : | | | : | | | : | | | :
Db      176 DDTSMYDNELWAPASEGLKPGPED---GPGKEAPELDEAEILDYLMDEVL   222
```

RESULT 12
US-10-076-069-2

```

; Sequence 2, Application US/10076069
; Publication No. US20020177214A1
; GENERAL INFORMATION:
; APPLICANT: JURECIC, ROLAND
; APPLICANT: NACHTMAN, RONALD
; TITLE OF INVENTION: HEPF, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL DEVELOPMENT
; FILE REFERENCE: 39532-176599
; CURRENT APPLICATION NUMBER: US/10/076,069
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/268,923
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;
; LENGTH: 237
;
; TYPE: PRT
;
; ORGANISM: Mus musculus
;
US-10-076-069-2

```

Query Match	13.2%;	Score 136;	DB 4;	Length 237;
Best Local Similarity	26.1%;	Pred. No. 0.00049;		
Matches 63;	Conservative 32;	Mismatches 86;	Indels 60;	Gaps 9;

```

QY      4 GLKRKHSDLBEEERKEWEMSPA-GLQSYQOALIRISLDKVQKSLGRAPSLRHHVLIHNTL 62
      |||||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
Db      5 GLKRKYGDQOEGVEGFVTVPSSYLQ--RQSLDMSLVKYLQCHMLVEPNLCRSVLIANTV 62

QY      63 QOLQAAALR-----LAP-APALPPEPLTGEEDFSLSATIGSILRELDTSMDGTEPPQ 113
      :|:|:|:| :|||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
Db      63 RQIQEEMSQDGVWGHGMAFQNVDRAPVERLVSTE---ILCRTVRGAEEHHPAPELEDAPLQ 119

QY      114 NPVTPLGLQNEVPQDPD-----VFLEALSSRYLGDGSLDDPF 151
      |||:|:| :|:|:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
Db      120 NSVSELPITVGSAPGQGNPNQSSLWEMDSPQENRGSFQKSLDQIFETLENK--NSSVEELF 177

QY      152 LDIDTSAVEKEPAR-----APPEPPHNLFCAPGSWEMNELDHIMEII 193
      ||:|:|:| :|:|:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
Db      178 SDVDSSTYYDLDTVLTCGMMSGTKSSLCNGLEGFAATPPPSSTCKS---DLAELDHVEIL 234

QY      194 L 194
      :

```

Db 235 V 235

RESULT 13
US-10-170

; Sequence 31, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:

```

; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 53682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-170-385-31

```

Query Match	12.1%;	Score 125;	DB 4;	Length 314;
Best Local Similarity	22.7%;	Pred. No. 0.007;		
Matches	71;	Conservative	34;	Mismatches 88;
				Indels 120;
				Gaps 12;

[illegible]

RESULT 14

US-10-755-889-627
 ; Sequence 627, Application US/10755889
 ; Publication No. US20040171823A1
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Bristol-Myers Squibb Company
 ;
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF- κ B
 ;
 ; TITLE OF INVENTION: PATHWAY
 ;
 ; FILE REFERENCE: D0284 NP
 ;
 ; CURRENT APPLICATION NUMBER: US/10/755, 889
 ;

; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 627
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-755-889-627

Query Match 12.1%; Score 125; DB 4; Length 314;
Best Local Similarity 22.7%; Pred. No. 0.007;
Matches 71; Conservative 34; Mismatches 88; Indels 120; Gaps 12;

QY 2 EGGLEKSHSDLEEEERWESPAGLOS-----YQALLRISLDKVQSLGPRAPSLRRH 55
DB 4 KGG-KRKFEHEDGLEGKIVSPCDGSPKSVYTLQRTIFNISLMKLYNHRPLTEPSLQKT 62
QY 56 VLIHNTLQOLAALR-----PALPPE---PLFLGEBDFSLATIGSLRELD 70
DB 63 VLIINMLRRIQELKQEGSLRPMFTPSQPTTEPSDSYREAPPAFSLHASPSSHPCDLGS 122
QY 71 -----LAPA-----PALPPE---PLFLGEBDFSLATIGSLRELD 103
DB 123 TTPLEACLTPTASLEDDDDTCTSQAMQPTAPTKLSPALLPEKD-SFSSALDEIHELCP 181
QY 104 TSMD-----GTEPPQNPTPLGLQNEVPPQ---PDPVFLAL-----SSRYLGDS 145
DB 182 TSTTEATAATDSVKGTSSEAGTQKLDGPQESRADDSKLMDSLPGNFEITTSIGFLTDL 241
QY 146 GLDD-FFLDIDTSAVEKEP-----ARAPPEPHNLF-----CAPGSWEW 183
DB 242 TLDDILFADIDTSMYDFDPTCTSSSGTASKMAPVSADLLKTLAPYSSQPTPSQDPFKMDL 301
QY 184 NELDHIMEITLGS 196
DB 302 TELDHIMEVLVGS 314

RESULT 15
US-11-095-870-12
; Sequence 12, Application US/11095870
; Publication No. US20050222034A1
; GENERAL INFORMATION:
; APPLICANT: Hsu, Stephen I-Hong
; TITLE OF INVENTION: MODULATION OF TRIP-BR FUNCTION AND METHOD OF TREATING
; TITLE OF INVENTION: PROLIFERATIVE DISORDERS
; FILE REFERENCE: 93231-89
; CURRENT APPLICATION NUMBER: US/11/095,870
; PRIOR FILING DATE: 2005-03-31
; PRIOR APPLICATION NUMBER: US 60/557,697
; PRIOR FILING DATE: 2004-03-31
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: TRIP-Br2
; US-11-095-870-12

Query Match 12.1%; Score 125; DB 6; Length 314;
Best Local Similarity 22.7%; Pred. No. 0.007;
Matches 71; Conservative 34; Mismatches 88; Indels 120; Gaps 12;

QY 2 EGGLEKSHSDLEEEERWESPAGLOS-----YQALLRISLDKVQSLGPRAPSLRRH 55
DB 4 KGG-KRKFEHEDGLEGKIVSPCDGSPKSVYTLQRTIFNISLMKLYNHRPLTEPSLQKT 62

QY 56 VLIHNTLQOLAALR-----PALPPE---PLFLGEBDFSLATIGSLRELD 70
DB 63 VLIINMLRRIQELKQEGSLRPMFTPSQPTTEPSDSYREAPPAFSLHASPSSHPCDLGS 122
QY 71 -----LAPA-----PALPPE---PLFLGEBDFSLATIGSLRELD 103
DB 123 TTPLEACLTPTASLEDDDDTCTSQAMQPTAPTKLSPALLPEKD-SFSSALDEIHELCP 181
QY 104 TSMD-----GTEPPQNPTPLGLQNEVPPQ---PDPVFLAL-----SSRYLGDS 145
DB 182 TSTTEATAATDSVKGTSSEAGTQKLDGPQESRADDSKLMDSLPGNFEITTSIGFLTDL 241
QY 146 GLDD-FFLDIDTSAVEKEP-----ARAPPEPHNLF-----CAPGSWEW 183
DB 242 TLDDILFADIDTSMYDFDPTCTSSSGTASKMAPVSADLLKTLAPYSSQPTPSQDPFKMDL 301
QY 184 NELDHIMEITLGS 196
DB 302 TELDHIMEVLVGS 314

Search completed: February 6, 2006, 10:16:34
Job time : 172 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 10:13:51 ; Search time 17 Seconds
(without alignments)
135.102 Million cell updates/sec

Title: US-10-069-386A-2
Perfect score: 1033
Sequence: 1 MEGGIKRKHSDLSEERWE.....APGSWEWNELDHIMEIILGS 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216	20.9	236	5	US-09-978-360A-568 Sequence 568, App
2	94	9.1	372	6	US-10-714-887-114 Sequence 114, App
3	90.5	8.8	324	6	US-10-511-314-19 Sequence 19, Appl
4	90.5	8.8	324	6	US-10-511-722-19 Sequence 19, Appl
5	88.5	8.6	1122	6	US-10-821-234-1657 Sequence 1657, Ap
6	88.5	8.6	1122	6	US-10-964-313-4 Sequence 4, Appli
7	84	8.1	1015	7	US-11-169-041-217 Sequence 217, App
8	83.5	8.1	1823	6	US-10-995-561-988 Sequence 988, App
9	83.5	8.1	2102	6	US-10-995-561-990 Sequence 990, App
10	83.5	8.1	2108	6	US-10-995-561-989 Sequence 989, App
11	83.5	8.1	2157	6	US-10-995-561-991 Sequence 991, App
12	82	7.9	2101	6	US-10-857-780-23 Sequence 23, Appl
13	81.5	7.9	384	7	US-11-134-563-2 Sequence 2, Appli
14	81	7.8	891	7	US-11-205-109-18 Sequence 18, Appl
15	80.5	7.8	606	6	US-10-055-877-58 Sequence 58, Appl
16	80.5	7.8	1532	6	US-10-821-234-914 Sequence 914, App
17	80.5	7.8	7968	7	US-11-143-980-49 Sequence 49, Appl
18	79	7.6	625	7	US-11-143-984A-39 Sequence 39, Appl
19	79	7.6	625	7	US-11-143-984A-110 Sequence 110, App
20	79	7.6	718	6	US-10-918-857-2 Sequence 2, Appli
21	79	7.6	790	6	US-10-918-857-6 Sequence 6, Appli
22	78.5	7.6	261	7	US-11-169-041-190 Sequence 190, App
23	78.5	7.6	1377	6	US-10-821-234-1070 Sequence 1070, Ap
24	77	7.5	558	6	US-10-504-364-3 Sequence 3, Appli
25	77	7.5	558	6	US-10-504-364-4 Sequence 4, Appli

26	77	7.5	1560	7	US-11-059-982-1	Sequence 1, Appli
27	76.5	7.4	616	6	US-10-982-545-5	Sequence 5, Appli
28	76.5	7.4	753	7	US-11-037-243-68	Sequence 68, Appl
29	76.5	7.4	1041	6	US-10-995-561-780	Sequence 780, App
30	76.5	7.4	1041	6	US-10-995-561-782	Sequence 782, App
31	76.5	7.4	1097	6	US-10-995-561-781	Sequence 781, App
32	76	7.4	863	7	US-11-169-041-167	Sequence 167, App
33	75.5	7.3	304	7	US-11-134-563-4	Sequence 4, Appli
34	75	7.3	559	7	US-11-150-845-38	Sequence 38, Appl
35	75	7.3	559	7	US-11-149-945-3	Sequence 3, Appli
36	75	7.3	672	7	US-11-000-463-455	Sequence 455, App
37	74.5	7.2	571	7	US-11-121-438-12	Sequence 12, Appl
38	74.5	7.2	884	6	US-10-995-561-786	Sequence 786, App
39	74.5	7.2	2630	7	US-11-186-731-2	Sequence 2, Appli
40	74.5	7.2	7968	7	US-11-186-731-5	Sequence 5, Appli
41	74	7.2	578	7	US-11-037-243-100	Sequence 100, App
42	74	7.2	746	6	US-10-828-831-5	Sequence 5, Appli
43	74	7.2	746	6	US-10-828-831-7	Sequence 7, Appli
44	74	7.2	1041	6	US-10-828-831-9	Sequence 9, Appli
45	74	7.2	1085	6	US-10-523-477-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-978-360A-568
; Sequence 568, Application US/09978360A
; Publication No. US20060009633A9
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclet, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 568
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -31..-1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28, 30, 40, 67, 86, 117, 120)
; OTHER INFORMATION: unknown


```

; CURRENT FILING DATE: 2004-10-18
; PRIOR APPLICATION NUMBER: 149217
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: 152183
; PRIOR FILING DATE: 2002-10-08
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-722-19

```

Query Match	8.8%;	Score 90.5;	DB 6;	Length 324;
Best Local Similarity	25.8%;	Pred. No. 0.24;		
Matches 54;	Conservative 17;	Mismatches 87;	Indels 51;	Gaps 10;

```

QY      3 GGLKRKHSDLBEEBERWESWPAGLOSXOALLRISLDKVQSRSLGPPRAPSILRRHVLINHTL 62
      |||| | : | : ||| : | |||| |
DB     36 GGLK---SPWRGEYKXBRHPPPNQANYHQ-----TLHAQPRELSPRAPGPRPAETTGRA 87
      |||| | | ||| : ||| : ||| :
QY     63 QQLQAALRLAPAPALPPE-----PLPLGEEDFSLATIGSILRELDTSMGTPEQPONPVT 117
      :|| | | ||| | | : | : ||| :
DB     88 PKLQ-----PPLPPEPPEPNKSPPLTSLKEB-----SGMWBPLPLSLBPAARNPSS 135
      |||| | | ||| : ||| : ||| :
QY    118 PLGLQNEVPRQ-----PDPVFLFALSSRY-----LGDGSLDDFFLDIDTSAVEKEPA 164
      | : ||| : ||| : ||| : | : ||| :
DB    136 P-ERKATVPEQELQQLLELFLNLSLQPSFLSEQEQILSLCLSTDSLSDDS---EKNPSS 191
      |||| | | ||| : ||| : ||| :
QY    165 RAPPPEPPHNLFCAPGSW-----EWN 184
      :|| | | ||| | | ||| :
DB    192 KASQSSRDTLSSGVHSSWSQAEARSSSWN 220
      |||| | | ||| : ||| : ||| :

```

```

RESULT 5
US-10-821-234-1657
; Sequence 1657, Application US/10821234
; Publication No. US20050255114A1
GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1657
; LENGTH: 1122
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1657

```

Query Match	8.6%;	Score 88.5;	DB 6;	Length 1122;
Best Local Similarity	24.4%;	Pred. No. 1.7;		
Matches	50;	Conservative	23;	Mismatches 75; Indels 57; Gaps 9;
QY	13	EEEEERWWS	PAGLS	YQ-----ALIRISLDKVRSLGPRAP---SL 52
		:::	:::	:::
Db	396	QGEAER-----	QALQSLRQGGTLTGKFMSTSSIPGCLLGVAB-----GDGSPFHGASL 444	
QY	53	RHVLINHTLQQLQALRLAPAPALPPEPLFLGEEDFSLSATIGSILRELDTSMDGTEPP 112		
		:::	:::	:::
Db	445	LQHVLL---LEQARQGSTLIAVPLHGQSPLYTGERVATSMRTVGKLPKRPRLSRQTSSP- 500		
QY	113	QNPVTPPLGLQNEVPPQPPDPVFLALSSR--YLGDGSLDFFLDIDITSAVEKEPARAPPEP 170		
		:::	:::	:::
Db	501	LPQSPQALQQLVMQQQHQQFLQKQKQQLQLGK-----ILTKGELPQPTTHPEET 552		

```

QY      171 PHNLFCAPGSWEWNELDHIMEIILG 195
          ||      ||: ||
DB      553 E-----EELTEQQEVLIG 565

```

```

RESULT 6
US-10-964-313-4
; Sequence 4, Application US/10964313
; Publication No. US20050287629A1
; GENERAL INFORMATION:
; APPLICANT: GROZINGER, CHRISTINA M.
; APPLICANT: HASSIG, CHRISTIAN A.
; APPLICANT: SCHREIBER, STUART L.
; TITLE OF INVENTION: CLASS II HUMAN HISTONE DEACETYLASES, AND USES RELATED
; TITLE OF INVENTION: THERETO
; FILE REFERENCE: HUV-037.02
; CURRENT APPLICATION NUMBER: US/10/964,313
; CURRENT FILING DATE: 2004-10-13
; PRIOR APPLICATION NUMBER: 09/800,187
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,802
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 4
; LENGTH: 1122
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-964-313-4

```

Query Match	8.6%;	Score 88.5;	DB 6;	length 1122;
Best Local Similarity	24.4%;	Pred. No. 1.7;		
Matches	50;	Conservative	23;	Mismatches 75; Indels 57; Gaps 9;
QY	13	EEBERWWS	PAGLOS	YQ-----ALLRISLDKVRSLGPAP-----SL 52
		:::	::	::
Db	396	QQEAE	R-----QALQSLRGGT	LTKGFMSTSSIPGILGVALE-----GDGSPHGASL 444
QY	53	RRHVLHINTLQQLAALRLAPAPALPEPL	FLGEEDFSLSATIGSILRELD	TSMDGTEPP 112
		:::	:::	:::
Db	445	LQHVLL---	LEQARQOSTLIAVPLHGOSPLVTGERVATSMRTVGKLP	RHRPLRSTQSSP- 500
QY	113	QNPVTPLGLQNEVPPQDPVFLEALSSR--	YLGDSGLDDFFL	LDIDTSAVEKEPARAPPEP 170
		::		:::
Db	501	-LPSPQALQQLVMQOQHQQFLEKQKQQLDLGK-----	ILTKGELPRQPI	THPEET 552
QY	171	PHNLFCAPGSGWENELDHIMEITLG	195	
			:::	
Db	553	E-----BELTEQQEVLIG	565	

```

RESULT 7
US-11-169-041-217
; Sequence 217, Application US/11169041
; Publication No. US20060019284A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
; TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE
; TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 10001 NP
; CURRENT APPLICATION NUMBER: US/11/169,041
; CURRENT FILING DATE: 2005-06-28
; PRIOR APPLICATION NUMBER: 60/584,405
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 527
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 217
; LENGTH: 1015
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-169-041-217

```


	Query Match	8.1%; Score 84; DB 7; Length 1015;
	Best Local Similarity	25.7%; Pred. No. 3.9;
	Matches	53; Conservative 20; Mismatches 63; Indels 70; Gaps 10;
QY	18 RWEWSPAQLQSYYOALRLRIS-----LDKVGRSI-----GPRAPS- 51 :: :: :: :: :	:
Dd	71 RYEVSFVALQRLRVALLXSGTGFTWQDDYTQYWMDQELADLPKTYLRRPEASSPARPSK 130 :: :: :: :: :	:
QY	52 -----LRRHVLIHNTLQQLQAALRIAPAPALPPEPLFLGEEDFSL 91 :: :: :: :: :	:
Dd	131 HSVGSERYSREGAALANALRRHLPLEBALSQAPASDVLAETHTAQDRPPABGDGRFS- 189 :: :: :: :: :	:
QY	92 SATIGSILRBLDTSMDGTEPPQNVTPLGLQNEVP-----QPDPVFLEALSSRYLGDS 145 :: :: :: :: :	:: :
Dd	190 ----ESILTVVAHTSATLTYP-GPRTQ--LRDDLPRTLGQLQPD-----ELSPKV--DS 235 :: :: :: :: :	:: :
QY	146 GLDDFFLDIDITSAVEKEPARAPEPP 171 :: :: :: :: :	:
Dd	236 GVDRHHL---MAALSAYAAQRPPAPP 258 :: :: :: :: :	:

```

RESULT 8
US-10-995-561-988
/ Sequence 988, Application US/10995561
/ Publication No. US20050272054A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
/ TITLE OF INVENTION: DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001559
/ CURRENT APPLICATION NUMBER: US/10/995,561
/ CURRENT FILING DATE: 2004-11-24
/ NUMBER OF SEQ ID NOS: 85702
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 988
/ LENGTH: 1823
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-995-561-988

```

Query Match	8.1%;	Score 83.5;	DB 6;	Length 1823;
Best Local Similarity	20.7%;	Pred. No. 8.7;		
Matches	45;	Conservative 18;	Mismatches 63;	Indels 91; Gaps 9;
QY	40	KVQRSLGPRAPSLRRHVL	-----HNTLQQLQALRLAPAPALPPEBFLGSEEDFSLSAT	94
Db	304	KVQKSLPRPRFORQQQEQLTKQQQQHQHQHQGGSAPPTPVPPSPQGV	-----T	352
QY	95	IGSI-----LRRLDTSMDGTEPQNPVT--PLGLQNEVPQEDVEFLG	-----	135
Db	353	LGAVPAPQAPPPPPKALYPGALGRPPPPMNFDPRRMM--IPPYVDPRLLGRRPLDFY		410
QY	136	-----ALSSRYLGDSG-----LDDFF-----		151
Db	411	PPGVHPGSLVPRERSDSGSSSEPFDRHAPAMLRERGTPPVDPKLANVGDVFTATPAEPR		470
QY	152	-----LDIDTSAVEKEPARAPPEPPHNLFCAPG		179
Db	471	PLTSPLRQADEDGKMRSETPPVPPPPY-LASYPG		506

RESULT 9
US-10-995-561-990
; Sequence 990, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559

```

; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 990
; LENGTH: 2102
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-995-561-990

```

Query Match	8.14;	Score 83.5;	DB 6;	Length 2102;
Best Local Similarity	20.74;	Pred. No. 10;		
Matches 45;	Conservative 18;	Mismatches 63;	Indels 91;	Gaps 9;

```
QY      40 KYQRSLGPRAPSLRRHVL-----HTLQQQLQAALRLAARALPPEPLFLGEEDFSLSAT   94
       ||:|||||:::||||:|||||:|
Db     583 KYQKSLPRPFQRQQQEQLKKQQQHQMQQHQGGSAAPTTPVPSPPOPY-----T    631
                                     :|
QY      95 ICSI-----LRDLDTSMDETPPNQNVLT--PLGLQNVEVPPQPDPVFLE----- 135
       ::||:|::|:|||||::|:|:|:|:|
Db     632 LGAVPAFAQAPPKPALKYPGALGRPEPMRPMNFDPRMM-IPRYVDPRLLQGRRPLLDFY 689
                                     |
QY     136 -----ALSSRYLGDSG-----LDDEF-----          151
       -|-|||
Db     690 PPGVHPSGLVPRERSDSCGSSEPFDRHAPAMLRBGTTPVDPKLAWGVFTATPAEPR   749
               |||||
QY     152 -----LDIDTSAVEKEPARAPEPPHNLFCAFG 179
       -|-||:|:|:|:|
Db     750 PLTSPLRQADEDCKMRSETPPVPPPPTY-LASTPYG 785
```

```

RESULT 10
US-10-995-561-989
/ Sequence 989, Application US/10995561
/ Publication No. US20050272054A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
/ TITLE OF INVENTION: DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001559
/ CURRENT APPLICATION NUMBER: US/10/995,561
/ CURRENT FILING DATE: 2004-11-24
/ NUMBER OF SEQ ID NOS: 85702
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 989
/ LENGTH: 2108
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-995-561-989

```

	Query Match	8.1%;	Score 83.5;	DB 6;	length 2108;	
	Best Local Similarity	20.7%;	Pred. No. 10;			
	Matches	45;	Conservative	18;	Mismatches	63;
					Indels	91;
					Gaps	9;
Qy	40 KVGKSLGPAPSLRRHVL-----NTLQQLQAALRLAPADLPPEPLFLGEEDFSLSAT	94				
Dd	583 KYQKSLPFRQRQQOQLKKOOQHQQHQOGSAPPPTVPESPFPQV-----T	631				
Qy	95 IGSIL-----LRLELDTSMDGTEPPQNPTV--PLGLONEVBPQPDPVFLE-----	135				
Dd	632 LGAVPAQAAPPPPCKALYFGALGRPPMPRPMNFDPRMMW--IPYVDRLLLQGRPLDFY	689				
Qy	136 -----ALSSRYLGDSG-----LDDEF-----	151				
Dd	690 PPGVHPISGLVPRERSDGGSSSEPFDRHAPAMLRERGTTPVDPEKLAWGVDTATPAEPR	749				
Qy	152 -----LDIDTSAVEKEPARAPPEPPHNLFCAFG	179				
Dd	750 PLTSPLROADEDDKGMRSETPPVPPPPPY-LASTYG	785				

RESULT 11


```

US-10-995-561-991
; Sequence 991, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 991
; LENGTH: 2157
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-991

```

	Query Match	8.1%;	Score 83.5;	DB 6;	Length 2157;
	Best Local Similarity	20.7%;	Pred. No. 11,		
	Matches	Conservative	18;	Mismatches	63; Indels 91; Gaps 9
QY	40 KYQRSLGPRAPSLRRHVL-----HNTLOQLQAALRLAPAPALPEEPLFLGEDFSLSAT	: :	:	:	:
Db	638 KYQKSLPFRFORQQOELTKQQQQHQWQQHQGGSAPTPTVPVPPSPQPV-----T	: :	:	:	:
QY	95 IGSI-----LRELDTSMDGTEBPQNVT--PLGLQNEVPPQPDPVFLE-----	: :	:	:	:
Db	687 LGAVPAQAPRPPPKALYPGALGRPPPMVMNFDPRMM--IPRYVDPRLLQGRPLDFY	: :	:	:	:
QY	136 -----ALSSRYLGDSG-----LDDFF-----	:	:	:	:
Db	745 PPGVHPSGLVPRERSDGGSSSBPFDRHADAMLRERGTPPVDPKLAMVGDVFTATPAEPR	:	:	:	:
QY	152 -----LDIDTSAVEKEPARAPPEPPHNLFCAFG	:	:	:	:
Db	805 PLTSPLRQADEDDEDKMRSETTPVPPPPPY-LASYFG	:	:	:	:

```

RESULT 12
US-10-857-780-23
; Sequence 23, Application US/10857780
; Publication No. US20050272043A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, RICHARD B.
; APPLICANT: BRAUN, ANDREAS
; APPLICANT: KAMERER, STEFAN M.
; APPLICANT: NELSON, MATTHEW ROBERTS
; APPLICANT: RENELAND, RIKARD HENRY
; APPLICANT: HOYAL-WRIGHTSON, CAROLYN R.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: SEQ-4069-CP
; CURRENT APPLICATION NUMBER: US/10/857,780
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: 10/723,681
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: 60/490,234
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: 60/525,239
; PRIOR FILING DATE: 2003-11-25
; NUMBER OF SEQ ID NOS: 4962
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 2101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-857-780-23

```

Query Match	7.9%;	Score 82;	DB 6;	Length 2101;
Best Local Similarity	27.1%;	Pred. No. 14;		
Matches 45;	Conservative 21;	Mismatches 76;	Indels 24;	Gaps 7;

[illegible]

```

RESULT 13
US-11-134-563-2
; Sequence 2, Application US/11134563
; Publication No. US20050287569A1
; GENERAL INFORMATION:
; APPLICANT: Leong, John M.
; APPLICANT: Campellone, Kenneth G.
; TITLE OF INVENTION: ESPFV NUCLEIC ACIDS AND PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 07917-280001
; CURRENT APPLICATION NUMBER: US/11/134,563
; CURRENT FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US 60/573,600
; PRIOR FILING DATE: 2004-05-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-134-563-2

```

Query Match	7.9%;	Score 81.5;	DB 7;	Length 384;
Best Local Similarity	23.5%;	Pred. No. 2;		
Matches 38;	Conservative 24;	Mismatches 51;	Indels 49;	Gaps 8;

```
QY      42 QNSLGPAPSLRRHVLINHTLOOLAAALRIA---PAPALP-PEPLFLGEEFSLSATIG   96
       : || : || : || : || : || : || : || : || : || : || : || : ||
Db     73 ESSLHQQLPNVRQRLLIQHLAEHGIXKARSMAEHIPPAPNWPAPPPPQNQSRLPDVAQ  132
QY      97 SIKELDTSMDGTEPPQN-----PVTPGLQNEVP-PQRPDVFLEALSSRYLGD  144
       ::|| : || : || : || : || : || : || : || : || : || : || : ||
Db    133 RLVOHL--AEHGIQPARNMMAEHIPPAPNWPAPRLLPVQNEQSRPLPD--VAORLVQH LAE  187
QY     145 SGLDDEFLLDIDTSAVEKEPAR-----APPEPPHN  173
       || : || : || : || : || : || : || : || : || : || : || : ||
Db    188 HGI-----QPASMAEHIPPAFNWPA PPPVQN  215
```

```

RESULT 14
US-11-205-109-18
; Sequence 18, Application US/11205109
; Publication No. US20050287641A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Staffa, Alfredo
; TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
; FILE REFERENCE: 3002-2US
; CURRENT APPLICATION NUMBER: US/11/205,109
; CURRENT FILING DATE: 2005-08-17
; PRIOR APPLICATION NUMBER: US/09/976,059
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/239,924
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
;
LENGTH: 891

```

TYPE: PRT
ORGANISM: Actinoplanes sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(1)
OTHER INFORMATION: V is a non-standard initiator codon. It is expected that the big
OTHER INFORMATION: synthesized protein will have a formylmethionine residue at this
OTHER INFORMATION: position
US-11-205-109-18

Query Match 7.8%; Score 81; DB 7; Length 891;
Best Local Similarity 27.5%; Pred. No. 6.2;
Matches 38; Conservative 17; Mismatches 61; Indels 22; Gaps 7;

QY 41 VQSLGPRAPSLRRHVLIHNTLQQLQALRLAP---APALPPEP---LFLGEDFSLSA 93
DB 34 IERRLARMFWEHVAAR---PGDEALRRRRELRARFVPEPGARAVLAVADGSADL 90
QY 94 TIGSILRELD-TSMGTEPPQNPVTPPLGLQNEVPPQDPVFLLEALSSRYLGDSGLDFFL 152
DB 91 VLVARDRDLRDALIALARPER--APGRKPAEPDAPP---SAAPAWGLGDGSPDRWA 145
QY 153 DIDTSAVEKEPARAPPEP 170
DB 146 EL-----RVPARGPADP 157

RESULT 15

US-10-055-877-58
Sequence 58, Application US/10055877
Publication No. US20050288241A1
GENERAL INFORMATION:

APPLICANT: Decristofaro, Marc
APPLICANT: Padigaru, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchernev, Velizar
APPLICANT: Zhong, Mei
APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
APPLICANT: Spytek, Kimberly
APPLICANT: Ratelli, Luca
APPLICANT: Kekuda, Ramesh
APPLICANT: Guo, Xiaojia
APPLICANT: Zehusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Mezes, Peter
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Eissen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shimkets, Richard
APPLICANT: Gusev, Vladimir
APPLICANT: Vernet, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25

PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 58
LENGTH: 606
TYPE: PRT
ORGANISM: Homo sapiens
US-10-055-877-58

Query Match 7.8%; Score 80.5; DB 6; Length 606;
Best Local Similarity 22.5%; Pred. No. 4.3;
Matches 48; Conservative 13; Mismatches 61; Indels 91; Gaps 10;

QY 3 GGLKRKHSDLBEEB-----RWESPAQLQSYQOALLRISLDKYQSLG-- 46
DB 4 GGVHTSVPLSQEAPSQTRAQTASAAGTWGRAPRGAPP-----PLSEAQSSGCG 55
QY 47 -----PRAPSLRRHVLIHNTLQQLQALRLAPAPALPPEPLFGEEDFSLATIGSLR 100
DB 56 GTSRSRPRPQAPAGRLASLALLRLRG-----PPRAMVLSQEB----- 93
QY 101 ELDTSMGTEPPQNPVTPPLGLQNEVPPQDPVFLLEALSSRYLGDSGLDFFLDIDTSAVE 160
DB 94 --PDSARGTSEAQ-PLGPAPTGAAPPPGPGF-----SDSPRAAVE 130
QY 161 K-----EPARA---PPEPPHNLFCAPGSWEW 183
DB 131 KVEVELAGPATAEPPHPPPE-----GGWGW 157

Search completed: February 6, 2006, 10:16:56
Job time : 18 secs